A DIGITAL GAME-BASED LEARNING FRAMEWORK **FOR ENHANCING PRIMARY SCHOOL STUDENTS'** CRITICAL THINKING SKILLS

MAHMOOD HAITHAM HUSSEIN

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
UNIVERSITI MALAYA
KUALA LUMPUR

2021

A DIGITAL GAME-BASED LEARNING FRAMEWORK FOR ENHANCING PRIMARY SCHOOL STUDENTS' CRITICAL THINKING SKILLS

MAHMOOD HAITHAM HUSSEIN

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
UNIVERSITI MALAYA
KUALA LUMPUR

2021

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Mahmood Hiatham Hussein

Matric No: 17027527/1

Name of Degree: Doctor of Philosophy

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

A Digital Game-Based Learning Framework for Enhancing Primary School Students' Critical Thinking Skills

Field of Study: e-Learning

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 06/12/2021

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

A DIGITAL GAME-BASED LEARNING FRAMEWORK FOR ENHANCING PRIMARY SCHOOL STUDENTS' CRITICAL THINKING SKILLS

ABSTRACT

Critical thinking is an essential skill that helps people to achieve success in a 21st-century global market that favours individuals capable of thinking critically. However, more traditional teaching methods have fallen short in terms of enhancing students' critical thinking skills. Thus, researchers are exploiting recent technological developments to incorporate technology in education. Among various other technology-supported teaching methods, digital game-based learning (DGBL) has been identified as an effective approach to promoting students' interest in learning and improving their critical thinking. Despite this interest, however, research shows that there is a paucity of DGBL applications based on specific frameworks, especially in the area of critical thinking. This research addresses this issue by proposing a framework comprising 6 components – inquiry, communication, mystery, decision making, challenge, and rewards (ICMDCR) – that aims to improve primary school students' critical thinking. It was hypothesised that the ICMDCR framework would significantly improve students' critical thinking and science subject scores. The researcher conducted a quasi-experimental study with an experimental group ($N = 62$) and a control group ($N = 62$) to test these hypotheses and assess the effectiveness of the framework. Results from repeated-measures analysis of variance (ANOVA) and post-hoc Bonferroni tests revealed that only students in the experimental group who played Ecoship Endeavour significantly improved their critical thinking (pre-test = 12.45, post-test = 15.10) and their science subject scores (pre-test = 56.66, post-test = 77.50). In addition, results showed that students in the control group who were taught via the traditional method experienced a considerable decrease in their critical thinking scores (pre-test = 11.98, post-test = 11.15), while their science scores did not significantly improve (pre-test = 55.32, post-test = 57.68). This framework can assist researchers and game developers in designing DGBL applications that refine students'

critical thinking capabilities. Moreover, students can use this game to enhance their critical thinking and participate in an engaging learning environment.

Keywords: Digital Game-Based Learning, DGBL, Critical thinking, DGBL framework, Science learning.

RANGKA PEMBELAJARAN BERASASKAN PERMAINAN UNTUK MEMUPUK KEMAHIRAN PEMIKIRAN KRITIS DALAM MURID SEKOLAH RENDAH

ABSTRAK

Pemikiran kritis adalah kemahiran yang asas bagi mencapai kejayaan dalam pasaran global abad ini, dan individu yang mempunyai kemahiran tersebut jelasnya mendapat kelebihan bandingan. Namun begitu, kaedah-kaedah pengajaran sering gagal memupuk kemahiran ini dalam pelajar-pelajar. Oleh itu, penyelidik sedang mengambil kesempatan dari kemajuan teknologi terkini dan menggabungkannya dalam pendidikan. Antara kaedah pengajaran yang mendukung teknologi, pembelajaran berasaskan permainan atau digital game-based learning (DGBL) telah dikenal pasti sebagai salah satu pendekatan yang efektif bagi mempromosi kemahiran untuk berfikir dengan kritis. Meskipun minat ini, kaji selidik menunjuk kekurangan aplikasi DGBL untuk rangka-rangka spesifik, terutamanya bagi pemikiran kritis. Kajian ini berdepan dengan isu ini dengan mencadang rangka yang mengandungi 6 komponen – inkuiri, komunikasi, misteri, mengambil keputusan, cabaran, dan ganjaran (ICMDCR) – yang berhasrat untuk memperbaiki kemahiran pemikiran kritis dalam kalangan murid-murid sekolah rendah. Ia telah disimpulkan bahawa rangka ICMDCR dapat meningkatkan pemikiran kritis pelajar serta pemarkahan mata pelajaran sains. Dengan menggunakan uji kaji pura-pura, eksperimen ini merangkumi kumpulan percubaan *(N = 62)* dan kumpulan terkawal *(N = 62)* untuk memeriksa hipotesis-hipotesis tersebut serta keberkesanan rangka. Keputusan analisis varians (ANOVA) dari langkah berulangan dan ujian post-hoc Bonferroni hanya menunjuk pembaikan markah pemikiran kritis bagi pelajar dalam kumpulan percubaan yang bermain Ecoship Endeavour (pra-ujian = 12.45, pasca-ujian = 15.10), serta markah mata pelajaran sains mereka (pra-ujian = 56.66, pasca-ujian = 77.50). Selaing itu, keputusan menunjuk bahawa kumpulan terkawal yang diajar mengguna kaedah

tradisional mengalami kemerosotan markah pemikiran kritis (pra-ujian = 11.98, pasca-ujian 11.15), manakala markah mata pelajaran sains tidak menunjuk pembaikan yang signifikan. Oleh itu, selain keberkesanan permainan ini untuk mempertingkatkan pemikiran kritis di kalangan pelajar dalam persekitaran yang kondusif, ia juga disimpulkan bahawa rangka ini mampu membantu penyelidik dan juga pembangun permainan dalam mereka bentuk aplikasi DGBL yang mampu memperhalus kebolehan pelajar-pelajar untuk memikir dengan kritis.

Kata kunci: Pembelajaran berasaskan permainan digital, DGBL, Pemikiran kritis, Kerangka DGBL, Pembelajaran sains.

ACKNOWLEDGEMENTS

First and foremost, I wish to express my heartfelt gratitude to my supervisor without whom I would not have been able to complete this research. I am deeply thankful for all her support, suggestions and patience towards the end, and all staff of University of Malaya, Faculty of Computer Science and Information Technology.

I also wish to express my sincere thanks to Dr. Su Moon Ting for her guidance in my research, and all staff of the University of Malaya, Faculty of Computer Science and Information Technology.

My sincere thanks go to my family who supported me from the first day I started working on this research. Their support helped me to get through the most difficult times.

I also would like to thanks express my sincere thanks to all my friends and colleagues for their cheerfulness and fun. Thank you for the great moments that we have shared together.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Symbols and Abbreviations

3 Cs	:	Correct, Complete, and Consistent
ABA	:	Applied Behaviour Analysis
ADDIE	:	Analysis, Design, Development, Implementation, and Evaluation
ANOVA	:	Analysis of Variance
DGBL	:	Digital Game-Based Learning
Four Cs	:	Creativity, Communication, Critical thinking, and Collaboration
Is framework	:	Identity, Immersion, Interactivity, Increasing Complexity, Informed Teaching
ICCE	:	Inquiry, Communication, Construction, and Expression
ICMDCR	:	Inquiry, Communication, Mystery, Decision making, Challenge and Rewards framework
ICT	:	Information and Communication Technology
IT	:	Information Technology
M	:	Mean
MDA	:	Mechanics, Dynamics, and Aesthetics
MOE	:	Ministry of Education
OECD	:	Organization for Economic Co-operation and Development
P value	:	Probability value
PISA	:	Program for International Student Assessment
R value	:	Correlation coefficient
RCT	:	Randomised Controlled Trial
SD	:	Standard Deviation
Std.Error	:	Standard Error
TIMSS	:	Trends in International Mathematics and Science Study

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

This chapter introduces the concept of digital gaming and its use in enhancing primary school students' critical thinking skills. This section also presents the research background, research questions, objectives, hypotheses, scope, and significance of the study.

1.1 Background Study

Information and communication technology (ICT) has allowed the world to become hyper-connected and enabled people to gain unlimited access to a vast volume of information. It has significantly increased competition and collaboration among citizens to ensure success and productivity at the global level (Bellanca, 2010; Siddiquah & Salim, 2017). As a result, a new set of skills, denominated under the category of 21st-century skills, has gradually been considered essential for success in both an academic and an employment context (Greiff, Niepel, & Wüstenberg, 2015).

Among these skills, critical thinking has received particular research attention (Qian & Clark, 2016). Critical thinking enables an individual to think reflectively and analytically, solve problems effectively, and judge skilfully (Kong, 2014; 2015). Fostering students' critical thinking skill is synonymous with success in undergraduate education and facilitates students' growth into more active and informed citizens (Halpern, 2014).

Employers, policymakers, politicians, and parents in Malaysia have unanimously and repeatedly stressed the importance of critical thinking and called for schools to reform and modernise themselves to prepare students for the challenges and demands of the 21st century (Fadhlullah & Ahmad, 2017; Rashid & Hashim, 2008; Zabit, 2010). Indeed, emphasis on the significance of critical thinking was established in the Malaysian Education Act (1996), which aimed to provide

> "an educational programme that includes curriculum and co-curricular activities which encompasses all the knowledge, skills, norms, values, cultural elements and beliefs to help develop a pupil fully with respect to the physical, spiritual, mental and emotional aspects as well as to inculcate and develop desirable moral values and to transmit knowledge". (Malaysian Education Act, 1996)

Furthermore, the Pelan Induk Pembangunan Pendidikan re-established the essentiality of critical thinking (2006–2010), noting that

> "Education plays an important role in developing human capital with a strong identity, competence, positive attitude, knowledgeable and high-skilled in order to fulfil the needs of the developed nation in 2020. The human capital to be cultivated should be able to think critically and creatively, to solve problems, having the capacity to create new opportunities, having the resilience and the ability to face the changing global environment". (Pelan Induk Pembangunan Pendidikan, 2006–2010)

Meanwhile, in preparation for Vision 2020, the development plan known as the Eleventh Malaysia called for a people-based and capital-based economy in the nation via the implementation of high-impact projects. This plan constituted six strategies, one of which entailed improving students' critical thinking: 'Teachers will embed Higher order thinking skills in their lessons to develop critical, creative, and innovative thinkers' (Economic Planning Unit, 2015, pp. 5–27).

Further, in a more detailed and comprehensive effort, the Malaysian Ministry of Education (MOE) introduced the National Education Blueprint (2013 – 2015). The proposed blueprint aimed to raise the standards of the Malaysian education to produce learners capable of advancing the nation to the forefront of the global economy and the market of the 21st-century (Salleh & Hatta, 2017). To achieve these goals and address the needs of the Malaysian people, the MOE stated that the educational system should place a greater emphasis on five aspirations, namely, access, quality, equity, unity, and efficiency. These aspirations are equally important, hence, no aspiration should be recognised as being more valuable or significant compared to other aspirations (Shan, Yunus, & Mohamad, 2018). Table 1.1 briefly describes these aspirations.

Table 1.1: MOE's Aspiration for the Educational System

Aspiration	Description
Access	The MOE aims to achieve full enrolment across all educational levels (e.g., from preschool to upper secondary school) by 2020, which will enable every Malaysian to reach their full potential (Kementerian Pendidikan Malaysia, 2013).
Quality	Secure a top third spot in large scale international assessments such as Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) within 15 years (Kementerian Pendidikan Malaysia, 2013).
Equity	Also in 2020, the MOE planned to halve the achievement gaps between students regardless of their gender, socioeconomics, and geographic (e.g., urban or rural) (Kementerian Pendidikan Malaysia, 2013).
Unity	The Malaysian MOE aspire to promote unity among Malaysians by encouraging them to interact with individuals from different socioeconomics, religions, and ethnic backgrounds. By doing so, students will understand, embrace and respect differences (Kementerian Pendidikan Malaysia, 2013).
Efficiency	To maximise students' learning outcomes within the budget allocated by the MOE (Kementerian Pendidikan Malaysia, 2013).

In the National Education Blueprint, the MOE provided a list of six key attributes, namely, knowledge, thinking skills, leadership skills, bilingual efficiency, ethics and spirituality, and national identity. According to the MOE, every student graduating from the Malaysian educational system should acquire these skills. Table 1.2 briefly describes these desired attributes.

Table 1.2: MOE's Aspiration for the Malaysian Student

Skill	Description
Knowledge	Every student should be literate, acquire mastery in essential subjects such as science and mathematics, and develop a general and well-rounded knowledge in Malaysia, Asia, and the world.
Thinking skills	According to the MOE, the education system in this area "historically fallen short, with students being less able than they should be in applying knowledge and thinking critically outside familiar academic contexts" (Kementerian Pendidikan Malaysia, 2013, p. 31). To address this issue, the MOE aims for every child to pursue a journey of lifelong learning and to synthesise and create new

	knowledge. By doing so, the MOE hopes that every child can develop a range of essential cognitive skills such as critical thinking, creativity, and innovations.
Leadership skills	The educational system will improve students' leadership skills by providing them with formal and informal learning environments that target four dimensions, namely, entrepreneurship, resilience, emotional intelligence, and strong communication skills (Kementerian Pendidikan Malaysia, 2013).
Bilingual efficiency	By promoting this attribute, the MOE hopes to equip every child with minimum proficiency in the Bahasa Melayu language, as this is the official language of the nation, and the English language, as English is the Lingua franca of the world.
Ethics and spirituality	The education system will inculcate students with strong ethics and spirituality to equip them with the knowledge required to face the challenges they are going to face in their adult life. By nurturing these values, the MOE hopes to produce caring individuals who contribute significantly to the betterment of Malaysia (Kementerian Pendidikan Malaysia, 2013; Salleh & Hatta, 2017).
National identity	Every student will identify as Malaysian regardless of their ethnic background, religion, and socioeconomics.

Despite repeated efforts to the contrary, the critical thinking competencies of undergraduate learners in Malaysia are not at the same level as those of other students worldwide. Scholars such as Rashid and Hashim (2008), Shaharom (2004), and Yunus et al. (2005) have investigated the issue and reached a similar conclusion: that is, the critical thinking skills of Malaysian undergraduates are considerably lower than that of their American counterparts. These observations suggest that existing schooling practices are not conducive to critical thinking. Such practices place significant emphasis on teaching students how to pass tests and obtain high scores in their examinations. These teaching methods have narrowed the scope of instruction and alienated students whose academic interests and strengths lie in other areas that are not commonly evaluated (Ehren & Hatch, 2013; Volante, 2004). Finally, this practice of 'teaching to the test' has another adverse effect: it provides students, particularly those heading to universities and colleges, with a false sense of security and academic aptitude (Ehren & Hatch, 2013; Smith & Fey, 2000; Volante, 2004). Consequently, these undergraduates could risk unemployment due to their poor critical thinking skills (Ambigapathy & Aniswal, 2005; Ting, Marzuki, Chuah, Misieng, & Jerome, 2017), which would, in turn, compromise Malaysia's position in the international market (Hanapi & Nordin, 2014; Ismail, Saifuddin, & Noraini, 2017; Rashid & Hashim, 2008).

Owing to the abovementioned factors, researchers and educators in Malaysia and globally have consolidated their efforts to overcome the limitations of traditional teaching methods by looking for new approaches to learning. Over the years, the integration of technology in education has been recognised as an effective student-centred learning approach that could provide students with rich learning environments (Nikou & Aavakare, 2021; Parkman, Litz, & Gromik, 2018).

As technology has increasingly advanced and evolved, interest in revolutionising the learning process has grown exponentially. In 2002, a learning concept emerged as scholars and educators aimed to capitalise on the enormous appeal and popularity of video games by incorporating these games into education (Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005; Behnamnia, Kamsin, Ismail, & Hayati, 2020; Chang, Warden, Liang, & Lin, 2018; Chen, Yang, Huang, & Fu, 2019). Scholars have recently proposed Digital Game-Based Learning (DGBL) as an effective teaching and learning method that provides learners with an enjoyable, interactive, and challenging environment (Chu & Chang, 2014; Liao, Chen, & Shih, 2019). DGBL is characterised as a learner-centred training and learning activity based on the integration of video games into educational content to achieve academic goals (Chen, Shih, & Law, 2020; Hwang, Sung, Hung, Huang, & Tsai, 2012; Hwang, Yang, & Wang, 2013; Li, Hwang, Chen, & Lin, 2021).

Using computerised game-based learning for education dates back to more than four decades ago. In 1976, a researcher from the University of Illinois named Bonnie Anderson developed the board game WEST. WEST is considered the first computer game-based learning environment and was designed to foster students' learning motivation and encourage them to play against and beat the computer using four arithmetic operations (Hong et al., 2009). In the early stages, simulations and games were used synonymously (Chen, 2021; Horn, 1995). It was not until further down the research road that a clear distinction was established (Sauvé, Renaud, Kaufman, & Marquis, 2007). Although this research discipline is decades old, it only began to receive attention from the research community and frequent utilisation in educational contexts in 2006 (Hwang & Wu, 2012).

1.2 Technology and Critical Thinking

According to Deana Kuhn, psychologist and author of the landmark book, *The Skill of Argument* (1991), critical thinkers are always in short supply. Kuhn found after studying hundreds of people from different walks of life that many of them were unable to demonstrate basic reasoning and argument skills (Gelder, 2001). Fortunately, scholars and educators have found that critical thinking is not an innate ability; rather, it constitutes a set of skills that can be acquired and improved via learning and training (Abrami et al., 2008; Ennis, 1989; Facione & Facione, 2008; Franco, Marques Vieira, & Tenreiro-Vieira, 2018).

Research concerning technology and critical thinking has observed that situating learners in a learning environment where rich interactions between the students and the learning material occur could lead to better thinking skills. For instance, in the last decade, the use of technology-rich classrooms has had a limited but constructive effect on students' thinking skills, particularly their evaluation capacities (Hopson, Simms, & Knezek, 2001). In addition to technology-rich learning environments, a study by McMahon (2009) perceived engagement as a significant factor in promoting students' critical thinking. Analysis of McMahon's experimental results indicated that the critical thinking of students who had more than five years of interaction with ICT was substantially higher than that of students who had less than five years of interaction. Moreover, a recent study by Wong and Cheung (2018) found that students' engagement with computers in education not only improved their critical thinking skills but also enhanced their creativity and problem-solving skills.

1.2.1 DGBL and Critical Thinking

Learning under the traditional method is passive: a teacher or an instructor essentially delivers the information with minimal student interventions (Hainey, Connolly, Stansfield, & Boyle, 2011; Lai, Hsiao, & Hsieh, 2018; Liu & Long, 2014). The traditional method is prevalent due to numerous factors. Among them is the fact that this method is suitable for delivering large volumes of data to students within a short time (Abdullah, Yaacob, Hashim, Hussain, & Roslan, 2019; Hackathorn, Solomon, Blankmeyer, Tennial, & Garczynski, 2011). In addition, scholars have noted other factors that impede students' ability to think critically. For example, students lack exposure to complex problems (Khalaf, 2018; Landsman & Gorski, 2007); there is an absence of engagement among students, and they might even drift to sleep (Licorish, Owen, Daniel, & George, 2018; Michel, Cater, & Varela, 2009).

When employing DGBL, students can participate in different playing conditions individually or collaboratively with teammates or via online forums (Hwang, Sung, Hung, Yang, & Huang, 2013; Sung & Hwang, 2013; 2018). Research surrounding DGBL has suggested that this learning method can help to improve students' critical thinking skills (Liao et al., 2019; Van Eck, 2006). While engaging in playing the game as a form of entertainment, students can simultaneously learn and solve problems (Van Eck, 2015). Researchers have also noted that students' interaction with DGBL applications could promote procedural knowledge, which is seldom taught through the traditional teaching method (Navarrete, 2014; Perini, Luglietti, Margoudi, Oliveira, & Taisch, 2018; Shaffer, 2006). Furthermore, learning via DGBL can increase students' interest in a topic (Chittaro, 2016; Harris, 2010), leading to students asking questions (Nicholson, 2014; Papanastasiou, Drigas, & Skianis, 2017) and enhancing students' research skills (Asbell-Clarke et al., 2012; Papanastasiou et al., 2017; Squire & Steinkuehler, 2005).

1.3 Problem Statement

Among various other technology-enhanced learning methods, DGBL has received increasing attention and been recognised as a promising learning method due to its potential to provide an authentic learning experience (Hwang et al., 2012, 2013, 2013a; Xu, Chen, Eutsler, Geng, & Kogut, 2019) and motivate learners of all ages (Chu & Chang, 2014; Hooshyar et al., 2020; Woo, 2014).

Despite growing interest in DGBL, however, research has shown that most DGBL applications are not developed based on specific learning frameworks or strategies (Behnamnia, Kamsin, & Ismail, 2020; Li & Tsai, 2013; Young et al., 2012). Scholars have also noted that DGBL applications have predominantly been used in an evaluative capacity to explore whether they could improve students' learning achievements. They are rarely implemented to promote students' 21st-century skills, such as critical thinking (Boyle et al., 2016; Clark, Tanner-Smith, & Killingsworth, 2016; Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Hainey, Connolly, Boyle, Wilson, & Razak, 2016).

According to the Partnership for 21st Century Skills, the 21st century skills includes communication, collaboration, creativity, and critical thinking or the 'four Cs' (Partnership for 21st-Century Skills, 2019). According to Claro et al. (2018), communication skills are based on recognising social norms to communicate information purposively for a specific audience. Ahmad (2020) and Fadli and Irwanto (2020) argued that communication in the 21st century is a complex skill that rooted is rooted not only in listening and speaking but also in showing empathy to those around you.

As for the second skill, creativity, this skill refers to individuals' capabilities to produce new and useful ideas or handle familiar issues or situation in a new manner, then utilise these ideas to create a product or service that is beneficial and potentially novel (van Laar, van Deursen, van Dijk & de Haan, 2017; 2020).

Concerning the third skill, collaboration, this skill refers to individual abilities to exchange ideas, argue, and disagree/agree on decisions with peers and supervisors whether in the digital or physical realms (Claro et al., 2018; van Laar et al., 2020). Finally, with regards to critical thinking, a detailed explanation of this skill is available in Chapter 2, Section 2.3.

Therefore, there is a need to develop a DGBL framework based on theoretically grounded concepts to improve students' critical thinking skills at the primary level. This study uses this framework to guide the design and development of a DGBL application, which will then be tested to examine the proposed framework's effectiveness.

1.4 Research Objectives

The objectives of this study are as follows:

- to develop a DGBL framework to enhance primary students' critical thinking skills;
- to develop a DGBL application that embeds the components of the proposed framework; and
- to evaluate the effectiveness of the proposed framework in enhancing the critical thinking skills of primary students in the learning of science.

1.5 Research Questions

This research aims to address the following questions, which relate to the objectives listed above.

- What are the suitable components required when designing a DGBL framework to improve critical thinking?
- How to develop a DGBL application using the proposed framework?
- What are the effects of a DGBL application (developed based on the proposed DGBL framework) on students' critical thinking skills and knowledge acquisition?

1.6 Research Hypotheses

The following hypotheses are established in this research.

H_0: The gaming application which is developed based on the proposed DGBL framework does not enhance students' critical thinking skills.

H_{1a}: The gaming application which is developed based on the proposed DGBL framework does enhance students' critical thinking skills.

$H2_0$: The gaming application which is developed based on the proposed DGBL framework does not enhance students' knowledge acquisition.

H_{2a}: The gaming application which is developed based on the proposed DGBL framework does enhance students' knowledge acquisition.

1.7 Research Scope

A quasi-experimental research study was conducted in a Malaysian public elementary school located in the state of Selangor in the Kajang district to cover the boundaries of the research problem, meet its objectives, and answer its questions. Despite the potential implications of this research, it is subject to a limited scope. Firstly, during childhood, students learn and acquire new skills effortlessly, hence, it is essential to equip and teach young learners such skills at an early age to help them progress and demonstrate better academic achievements in higher classes such as secondary schools (Murugan & Rajoo, 2013; Wu & Lin, 2016). As such, this research only targeted primary-aged students.

Secondly, DGBL applications could be utilised to promote various 21st-century skills, such as creativity or complex problem-solving, or other types of cognitive skills, such as computational or logical thinking. As such, the scope of this research is to examine the impact of the DGBL approach in improving primary-aged students' critical thinking via the lens of the proposed DGBL framework.

Thirdly, the viewpoints of teachers have been considered at various stages throughout the development process. However, the study reviewed their perspectives purely in relation to developing a more reliable and effective gaming application. As such, their attitudes and perceptions regarding DGBL or how this method of instruction would affect their critical thinking skills have not been tested or evaluated in this research.

Finally, the study scope solely comprises DGBL; hence, other technology-enhanced learning approaches or their role in improving learners' critical thinking have not been examined.

1.8 Research Significance

A review of the existing literature pertaining to the integration of computer games in education indicates that researchers have recognised DGBL as a promising learning approach (Gee, 2003; Hung, Yang, Hwang, Chu, & Wang, 2018; Ke, 2016). Indeed, this research area has received growing scholarly attention (Boyle et al., 2016; Connolly et al., 2012; Hainey et al., 2016); however, a closer look at the literature reveals several issues and challenges.

In the DGBL domain, the majority of research works, regardless of their intended outcomes or the subject they are addressing, have been designed in a proof-of-concept manner. Such an approach focuses on simple questions, such as 'can games support learning?'. Thus, a limited number of studies have explored the effectiveness of the DGBL approach from the perspective of a framework based on specific learning principles and design strategies (Clark et al., 2016; Li & Tsai, 2013). While Behnamnia et al. (2020) noted that DGBL applications are being increasingly utilised in both schools and homes, the authors also observed that there are few frameworks or models that combine multiple elements with the goal of improving students' 21st-century skills.

Most DGBL applications have been developed to teach students low-level thinking skills, such as drill and practice (Westera, 2019), instead of focusing on students' cognitive skills (Clark et al., 2016; Fiorella & Mayer, 2012) or helping learners to develop a deep level of knowledge and understanding based on reliable learning principles and theories (Westera, 2019).

Further, according to Santos and Fraternali (2015), most studies that proposed DGBL frameworks identified empirical validation of their frameworks as an issue left for the future.

1.9 Thesis Outline

This section provides a brief description of how the following chapters in this dissertation are organised.

- Chapter 2 consists of a review of the prior research, focusing on two central themes. The first theme begins with an overview of critical thinking and its definitions. The chapter then addresses the significance of critical thinking, highlights the implications of poor critical thinking, and discusses the barriers to adopting critical thinking successfully and effectively in classroom settings. The second theme of this chapter provides an outline of the DGBL method and highlights a number of its definitions. The subsequent sections report on DGBL research in the domains of students' academic achievements in general, their achievement in the science subject, and critical thinking. Finally, the chapter details various DGBL frameworks and compares their strengths and weaknesses.

- Chapter 3 elaborates the various steps that went into the development of the gaming DGBL framework from interviewing teachers and students, identifying limitations in an existing DGBL framework, proposing the DGBL framework, develop a DGBL application based on the proposed framework, and finally elucidate how the components of the framework were integrated into the gaming application.
- Chapter 4 clarifies the evaluation of the prototype and highlights the results of the study.
- Chapter 5 compares the outcomes of this research to the results of previous research works, concludes the thesis and identifies the study's contribution to knowledge and its limitations and implications before offering recommendations for future studies.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter provides background information pertaining to critical thinking since it is essential to offer a historical overview of critical thinking, its definitions, and its importance for adults and young learners. This chapter highlights the ramifications of flawed critical thinking and the barriers and obstacles that inhibit the adequate integration of critical thinking in education, with a specific interest in the Malaysian educational system. This section also presents a detailed view of key research works in the area of DGBL and science education. This analysis of DGBL literature will help researchers and educators gain a clearer picture of this research area and recognise its potential, limitations, and challenges to ensure DGBL applications can be used to improve students' learning in general and their critical thinking in particular. Several relevant DGBL frameworks are evaluated to offer an insight into the mode of operation of these guidelines and frameworks and how to use them to design effective and reliable learning tools.

2.2 Overview of Critical Thinking

Teachers, educators, policymakers, and employers have always endeavoured to improve and enhance students' abilities to think critically. The literature shows that the academic roots of critical thinking go back 2,500 years: researchers and historians first attributed critical thinking to Socrates when he 'discovered a method of probing questioning that people could not rationally justify their confident claims to knowledge' (Paul, Elder, & Bartell, 1997, p. 4).

Socrates encouraged people to contemplate the common understandings that are seldom examined in the contemporary pursuit of relevance and truth. Furthermore, Socrates challenged the assumptions and beliefs of those in positions of power and founded Socratic questioning, which is a procedure of reflective and systematic thinking that advances the significance of raising deep questions and looking for proof to assess rhetoric. Socrates's quest for the essence of reason and truth promoted a detailed assessment of accounts and a comprehension of their evidence, assumptions, theories, reasoning, and implications (Heiniger, Clark, & Egan, 2018; McGuire, 2010). The Socratic method underpins the thinking of other philosophers in Ancient Greece, such as Aristotle and Plato, and has contributed to the works of other thinkers over the centuries, including Thomas Aquinas, Erasmus, Francis Bacon, Descartes, Sir Thomas Moore, Hobbes and Lock, Robert Boyle, and Sir Isaac Newton (Paul & Elder, 2007; Paul, Elder, & Bartell, 1997a).

The magnitude of critical thinking continued to persist throughout the 20th century; as a result, philosophers and educators have often debated this concept. One of the most notable contributions in this period was made by John Dewey. Dewey was a philosopher, psychologist, and educator from the United States who is often recognised as the 'father' of contemporary critical thinking (Fisher, 2011; McGuire, 2010). Dewey's pragmatic approach to critical thinking recommends a student-centred education and emphasises that teaching students critical thinking is the primary purpose of education and learning (Dewey, 1933; Hepner, 2012). Piaget's (1977) ideas concurred with those of Dewey. Piaget argued that learning is an active process, not a static experience. The more active students are in an activity, the more learning that occurs. When learning, students make mistakes, seek guidance and advice, modify actions, and gain knowledge that can be applied to real-life situations.

In the 21st century, the necessity of critical thinking was identified by the Partnership for 21st Century Skills, a partnership established by the U.S. Department of Education in 2002 accompanied by other major international business corporations, including Apple, Inc., Cisco Systems, Inc., the Microsoft Corporation, and AOL Time Warner Foundation, among others. The goal behind the creation of this organisation was to inaugurate a vision for learning and teaching in the 21st century and ensure success and prosperity to students as citizens of this new era (Eye, Gilb, & Hicks, 2013; Partnership for 21st Century Skills, 2009).

2.3 Definitions of Critical Thinking

Although many psychologists and educators have offered definitions of the term 'critical thinking', these definitions are similar in that they all emphasise that critical thinking is an intellectual and cognitive process that entails logic and reason. There are also, however, some differences in these definitions. In recent decades, critical thinking has become a desirable human trait, and teaching critical thinking in schools wherever possible has become the goal of many educators (Abduljabbar, 2019; McPeck, 1990). In order to understand how critical thinking influences the learning process, it is fundamental to look at how it has been classified by different educators and scholars.

Facione (1990) interpreted critical thinking as "purposeful, self-regulatory judgement, which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgement is based" (p. 3). While Ennis (2001) defined critical thinking as "reasonable reflective thinking focused on deciding what to believe or do" (p. 12), Paul (2005) argued that critical thinking is "the art of thinking about thinking in an intellectually disciplined manner" (p. 28).

According to Scriven and Paul (2007), meanwhile, critical thinking is "the intellectually disciplined process of actively and skilfully conceptualising, applying, analysing, synthesising, and/or evaluating information gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action" (p. 1). Willingham (2008) classified critical thinking as "seeing both sides of an issue, being open to new evidence that confirms your ideas, reasoning dispassionately, deducing and inferring conclusions from available facts and solving problems". Finally, Lai (2011) viewed critical thinking as the "skills of analysing arguments, making inferences using inductive or deductive reasoning, judging or evaluating, and making decisions or solving problems" (p. 2).

2.4 The Importance of Critical Thinking

In the information-oriented society in which we currently live, technology is ubiquitous, and the Internet can be accessed from virtually anywhere and at any time. This new era of access to unlimited amounts of information has triggered a recent change among students (Perry, 2014). Today, students have no problem locating the desired information. However, the issue is not about finding the desired information; what is of greater concern is validating the reliability of the information found (Perry, 2014; Tiruneh, De Cock, Weldeslassie, Elen, & Janssen, 2017). The skills required to succeed in this new era of excessive and often inconsistent information are known as 21st-century skills. Among the several skills deemed appropriate for a 21st-century education, critical thinking has received the most scholarly attention and research interest (Qian & Clark, 2016).

Numerous researchers have highlighted the importance of teaching students how to think critically, supported by decades of empirical and theoretical research (Lai, 2011; Zohar & Cohen, 2016). For example, Facione (1990) suggested that improving students' critical thinking capabilities should be a primary goal for every grade in the K–12 curriculum. Sternberg (2004) argued that instructing students how to think critically and analytically should be among the top priorities of current educational institutions. Sternberg (2004), and later on, Gelerstein, Nussbaum, Chiuminatto, and López (2016), claimed that critical thinking permits students to make independent decisions and to investigate a given phenomenon when this phenomenon is not based on robust evidence.

Moreover, McGuinness (2005) stated that critical thinking motivates people to use reason and logic to investigate a given issue or topic from multiple angles; thus, they can form their own thoughts and attitudes with confidence instead of merely regurgitating the anecdotes of other people. According to Bonney and Sternberg (2016), one of the most vital jobs for teachers in the classroom is to teach students how to learn and become critical thinkers rather than solely transferring knowledge. The children who are admitted to schools in 2018 will be young adults in 2030, which suggests that educational institutes, along with teachers, need to prepare these future employees for occupations that have not yet been created, for technologies that have not yet been invented, and to fix issues that have not yet been envisioned. To be successful in this uncertain period, these students should develop a set of competencies, including critical thinking skills, that will enable them to forecast what could be needed in the future or how decisions made today might influence future events (Organisation for Economic Co-operation and Development [OECD], 2018).

One of the main reasons why scholars and educators have called for the promotion of students' critical thinking is, essentially, economically motivated. Indeed, the emergence of a knowledge-based economy over a once-dominant manufacturing economy means that 21st-century jobs demand a highly educated workforce capable of higher order thinking skills to compete and perform in roles within a fiercely tight and competitive global market. Therefore, students with advanced critical thinking abilities and high standards of work ethics will be favoured for high-paying and sophisticated economic opportunities (Advisory Council on Economic Growth, 2017; Renzulli, 2008; Marx, 2011). In addition, to adapt to a dynamic and technology-laden employment market, students should be able to think analytically and critically, solve problems efficiently, and communicate effectively in the workplace (Kalonji, 2005, Kong, 2015). According to the Partnership for 21st Century Skills, the four Cs have been recognised as proficiencies that distinguish students who are willing to participate in inherently complicated living and working environments in today's world and those who are not (Partnership for 21st Century Skills, 2019).

2.4.1 The Importance of Critical Thinking for Young Learners

Many psychologists, researchers, and educators have indicated that people are not born with innate critical thinking abilities; rather, these skills are taught via adequate education and training (Abrami et al., 2008; Butler, 2012; Facione & Facione, 2008; Persky, Medina, & Castleberry, 2019), and can be developed and improved at an early age (Gelerstein et al., 2016; Lai, 2011; Willingham, 2008; Ugwuozor, Ede, Ifelunni, & Abiogu, 2020). Numerous studies underpin this assumption, indicating a reliable connection between critical thinking capacities and early training and education. For example, Ennis (1989) suggested that the most appropriate time to promote students' critical thinking capacities is during elementary education. Meanwhile, Koenig and Harris (2005) showed that children aged 3–4 years old are capable of discerning the authenticity of different sources of information. Heyman and Legare (2005) also observed that children aged 7–10 years old became noticeably mindful of the fact that certain individuals may have reasons to conceal the truth behind some information or events. In contrast, children below that age were not regularly critical of the reliability of people with such motives.

Furthermore, Pillow (2008) asserted that although a critical thinking curriculum is rarely taught for young learners, children are capable of critical thinking when communicating with their friends and adults. Pillow argued that children experience a transition in their cognitive development during the elementary years (when children are aged 4–5 and 6–7). Children evolve from believing what they see without question to grasping multiple interpretations of the same information. While Hwa (2016) argued that 21st-century skills are most visible in tasks performed by adult individuals, the scholar nonetheless maintained that the development of these skills begins at an earlier age and recognised primary education as the foundation for further cognitive skills development.

2.5 The Implications of Deficient Critical Thinking

If critical thinking skills are not well-developed through the educational system, there will be far-reaching ramifications for our ability to participate in social development and build a productive workforce (Flores, Matkin, Burbach, Quinn, & Harding, 2012; Gelerstein et al., 2016). Sternberg (2004) stated that although students may be knowledgeable, they may not have been taught how to think critically and analytically and, consequently, could become highly vulnerable to the fallacious reasoning exhibited by political leaders and within the media in its various forms.

Moreover, a study of over 400 U.S. employers demonstrated that business leaders are looking for job entrants to come out of schools with the fundamental and applied skills necessary for a 21st-century work environment. However, current high school graduates are entering the workforce with poor mastery of knowledge and insufficient skills required to accomplish prosperous careers (Casner-Lotto & Barrington, 2006; Lowther, Inan, Strahl, & Ross, 2012). Stewart, Wall, and Marciniec (2016) conducted a study of 214 college students to ascertain whether students felt confident in their soft skills, such as communication, critical thinking, and problem-solving. Their findings revealed that most students rated their soft skills very highly. Nevertheless, the skills in which students feel competent are the same skills that employers believe graduates are short of possessing.

2.5.1 The Implications of Deficient Critical Thinking in Malaysia

Malaysian leaders and policymakers aimed to make Malaysia a developed nation by 2020 (Ahrari, Samah, Hassan, Wahat, & Zaremohzzabieh, 2016; Salleh, 2006) by emphasising the importance of education (Fadzil & Saat, 2014) and producing a workforce capable of innovating and competing in the global market (Meng, Idris, & Eu, 2014). Despite these aspirations, students in Malaysia are still graduating high schools with worryingly low levels of critical thinking skills (Ismail, Harun, Zakaria, & Salleh, 2018; Siti Rahayah Ariffin & NorAzaheen, 2009). This has resulted in students' poor achievement in PISA and TIMSS.

PISA is a triennial international survey established by the OECD. It intends to assess numerous nations' educational systems by examining the knowledge and skills of 15-year-old students' performances in three areas: science, reading, and mathematics (OECD, 2018).

In 2009, Malaysia ranked 53 out of 74 countries in the science domain. Malaysian students scored an average of 422 points, which was lower than the OECD average of 501 points, a deficit of 79 points. There was a significant gap between Malaysia and Singapore, whose students scored an average of 542 (Abdullah & Peters, 2015). In 2012, on the science scale, Malaysian students lost two points from the previous test and scored an average of 420 points, which was below the OECD average of 501 points. Altogether, the students' performance ranked Malaysia at 52 out of 65 participating nations. The PISA test results revealed that the gap between Malaysia and Singapore increased to 131 points on the science scale (Abdullah & Peters, 2015). In 2015, a new domain was introduced to assess students' competencies in collaborative problem-solving (OECD, 2017). However, in this round of the PISA test, only half of the schools randomly chosen to participate actually took the assessment. As a result, Malaysia was excluded from the official database of the PISA results (Thomson, Bortoli, & Underwood, 2016).

In the last cycle of the PISA test, Malaysian students scored 438 points on the science scale, an increase of 18 points. However, despite this promising improvement in performance. Malaysia was 51 points below the OECD average of 489 points (Gurria, 2018). As such, Malaysia landed the 48th spot out of 78 participating nations. Further, the gap between Malaysia and Singapore on the science scale remained considerably large at 113 points. Furthermore, only 1% of Malaysian students were top performers in science, which was lower than the OECD average of 7%, and significantly lower than 21% scored by Singapore's top performers (OECD, 2019, 2019a). To provide a clearer picture pertaining to the performance of Malaysian students. Table 2.1 lists the findings of other Asian countries along with Malaysia in the three scales of the PISA test.

Table 2.3: Malaysia Performance in the Last Three PISA Cycles

			2012			
Country	**Mathematics**		**Reading**		**Science**	
	Points	**Position**	**Points**	**Position**	**Points**	**Position**
China	613	1	570	1	580	1
Singapore	573	2	542	3	551	3
Hong Kong	561	3	545	2	555	2
Taiwan	560	4	523	8	523	13
South Korea	554	5	536	5	538	7
Macau	538	6	509	16	521	17
Japan	536	7	538	4	547	4
Vietnam	511	17	508	19	528	8
Thailand	427	50	441	48	444	48
Malaysia	421	52	398	59	420	53
Indonesia	375	64	396	61	382	64

Results obtained from (Da Wan, Sirat & Razak, 2018)

			2015			
Country	**Mathematics**		**Reading**		**Science**	
	Points	**Position**	**Points**	**Position**	**Points**	**Position**
China	531	6	494	27	518	10
Singapore	564	1	535	1	556	1
Hong Kong	548	2	527	3	523	9

Taiwan	542	4	497	23	532	4
South Korea	524	7	517	7	516	11
Macau	544	3	509	10	529	6
Japan	532	5	516	8	538	2
Vietnam	495	22	487	30	525	8
Thailand	415	54	409	57	421	54
Malaysia	NA	NA	NA	NA	NA	NA
Indonesia	386	63	397	64	403	62

Results obtained from (OCED, 2016)

2018

Country	**Mathematics**		**Reading**		**Science**	
	Points	**Position**	**Points**	**Position**	**Points**	**Position**
China	591	1	555	1	590	1
Singapore	569	2	549	2	551	2
Hong Kong	551	4	524	4	517	9
Taiwan	531	5	503	16	517	10
South Korea	526	7	514	9	519	7
Macau	558	3	525	3	544	3
Japan	527	6	504	14	529	5
Vietnam	NA	NA	NA	NA	NA	NA
Thailand	419	57	393	66	426	54
Malaysia	440	47	415	56	438	48
Indonesia	379	71	371	72	396	70

Results obtained from (Schleicher, 2018)

TIMSS is an international comparative assessment administrated since 1995 and takes place every four years. This large-scale assessment aims to evaluate trends in mathematics and science at the fourth- and eighth-grade levels (Provasnik et al., 2016). However, Malaysian participation was limited to eighth-grade levels only. In addition to scholastic performances in science and mathematics, TIMSS also conducts surveys among the students, teachers, and principals to study the background of the participants such as home and school contexts in learning science and mathematics, curriculum implementation, instructional practices, and school resources (Phang, Khamis, Nawi, & Pusppanathan, 2020).

By observing how Malaysian students performed in this assessment, one could argue that students in Malaysia experienced two phases. The first phase occurred in the 1999 and 2003 cycles, when students scored 492 and 510 points, respectively. Both performances were above the TIMSS international average of 488 and 474 points, respectively. In addition, during this phase, specifically in 1999, 5% of Malaysian students performed at the advanced international benchmark, that is students scoring an average of 616 points or more. Additionally, in 2003, 4% of Malaysian students performed at the advanced international benchmark, that is students scoring an average of 625 points or more. Finally, in this phase, the performance gap between Malaysia and Singapore was 76 points in 1999, and then it was further reduced to 68 points in 2003 (Martin, Mullis, Foy, & Stanco, 2012; Ng, Lay, Areepattamannil, Treagust, & Chandrasegaran, 2012).

The second phase is marked with fluctuations and inconsistencies, it occurred throughout the last four TIMSS cycles from 2007 to 2019. In this phase, Malaysian students scored 471 points in 2007, 426 points in 2011, 471 points in 2015, and 460 points in 2019. As a consequence, not even once were Malaysian students able to surpass or reach the international average of 500 points. Moreover, during this phase, the number of students performing at the advanced international benchmark, that is students scoring an average of 625 points or more, fell down to 3% in 2007, 2015, and 2019, and only 1% in 2011. Further, during the last four cycles, Malaysian students were unable to reach or break the record of 510 points they sat more than 16 years ago in 2003. Furthermore, due to this inconsistent performance, the gap between Malaysia and Singapore remarkably increased to 96 points in 2007, 164 points in 2011, 126 points in 2015, and 148 points in 2019.

Table 2.2 provides a broader perspective with regards to the performance of Malaysia against other Asian nations in both the science and mathematics domain over the last four cycles of the TIMSS assessment.

Table 2.4: Malaysia Performance in the Last Four TIMSS Cycles

			2007			
Country	**Science score 8th grade**	**Rank**	**% of students at Advanced Benchmark**	**Mathematics score 8th grade**	**Rank**	**% of students at Advanced Benchmark**
Singapore	567	1	32	593	3	40
Hong Kong	530	9	10	572	4	31
Taiwan	567	2	25	598	1	45
South Korea	567	4	17	597	2	40
Japan	567	4	17	570	5	26
Thailand	471	21	3	441	29	3
Malaysia	471	21	3	474	20	2
Indonesia	427	34	0	397	36	0
Results obtained from (Mullis, 2008; Martin, Mullis, & Foy, 2007).						
			2011			
Country	**Science score 8th grade**	**Rank**	**% of students at Advanced Benchmark**	**Mathematics score 8th grade**	**Rank**	**% of students at Advanced Benchmark**
Singapore	590	1	40	611	2	48
Hong Kong	535	8	9	586	4	34
Taiwan	564	2	24	609	3	49
South Korea	560	3	20	613	1	47
Japan	558	4	18	570	5	27
Thailand	451	27	1	427	28	2
Malaysia	426	32	1	440	26	2
Indonesia	406	40	0	386	38	0
Results obtained from (Martin, Mullis, Foy, & Stanco, 2012; Mullis, Martin, Foy, & Arora, 2012).						
			2015			
Country	**Science score 8th grade**	**Rank**	**% of students at Advanced Benchmark**	**Mathematics score 8th grade**	**Rank**	**% of students at Advanced Benchmark**
Singapore	597	1	42	621	1	54
Hong Kong	546	6	12	594	4	37
Taiwan	569	3	27	599	3	44
South Korea	556	4	19	606	2	43
Japan	571	2	24	586	5	34
Thailand	456	28	-	431	30	3
Malaysia	471	24	3	465	22	3
Results obtained from (Mullis, Martin, Foy, & Hooper, 2016)						
			2019			

Country	Science score 8th grade	Rank	% of students at Advanced Benchmark	Mathematics score 8th grade	Rank	% of students at Advanced Benchmark
Singapore	608	1	48	616	1	51
Hong Kong	504	17	9	578	5	32
Taiwan	574	2	29	612	2	49
South Korea	561	4	22	607	3	45
Japan	570	3	22	594	4	37
Malaysia	460	29	3	461	28	4

Results obtained from (Mullis, Martin, Foy, Kelly, & Fishbein, 2020)

The TIMSS assessment also measures students' cognitive skills in the subject of science along three domains, namely, the knowing, applying, and reasoning domains. The knowing domain addresses the facts, concepts, and procedures students are required to know, acquiring a wide range of factual knowledge will significantly help learners face the more challenging cognitive tasks native to the scientific inquiry process, the applying domain is primarily concerned with students' abilities to put their knowledge and conceptual understanding into practice in order to solve problems or answer questions. Finally, the reasoning domain refers to going beyond the solution of familiar problems to encompass unfamiliar situations, complex contexts, and multistep problems (Mullis et al., 2016; 2020).

A closer examination of students' performance in these domains revealed that in 2015 Malaysian students achieved their highest score in the knowing domain (466 points), and applying domain (476 points). However, with regards to the reasoning domain, the most challenging of all cognitive domains, the highest score attained by Malaysian students was more than 14 years ago, in 2007, (487 points). In addition, over the last four cycles of this assessment, Malaysian students never reached or surpassed the 500 points in any cognitive domain targetted by the TIMSS (Gonzales et al., 2008; Mullis et al., 2016; 2020). Further, the unsatisfactory performance of Malaysian students in the cognitive domains of science was not a coincidence. Additional analyses of the TIMSS revealed that Malaysian students achieved their highest score in 2007, and in a similar manner t their performance in science, Malaysian students did not reach or surpass the 500 points in any cognitive domain addressed by the TIMSS.

By comparing the performance of students in Malaysia with their counterparts in Singapore a similar performance gap emerges. In Singapore, in science and mathematics cognitive domains, students always performed over 550 points and their performance consistently improved from one cycle to the next. Further, in the cognitive domains of science, students in Singapore broke the barrier of 600 points on three occasions two of which were in the latest cycle of TIMSS in 2019 (Galia, 2008; Martin et al., 2012; Mullis et al., 2016; 2020), whereas, in the cognitive domains of mathematics, Singaporean students broke the 600 points barrier in every domain and in every cycle since 2011.

Table 2.3 aims to provide a detailed look on the performance of Malaysian students in relation to other neighbouring Asian countries.

Table 2.5: Malaysia Performance in the Last Four TIMSS Cycles in the Cognitive Domains

			2007			
		Science			Mathematics	
Country	**Knowing**	**Applying**	**Reasoning**	**Knowing**	**Applying**	**Reasoning**
Singapore	554	567	564	581	593	579
Hong Kong	532	522	533	574	569	557
Taiwan	565	560	541	594	592	591
South Korea	543	547	558	596	595	579
Japan	534	555	560	560	565	568
Thailand	473	472	473	436	446	456
Malaysia	458	473	487	477	478	468
Indonesia	426	425	438	397	398	405
Results obtained from (Martin et al., 2007; Mullis & Martin, 2008)						
			2011			
		Science			**Mathematics**	
Country	**Knowing**	**Applying**	**Reasoning**	**Knowing**	**Applying**	**Reasoning**

Singapore	588	589	592	617	613	604
Hong Kong	544	529	538	591	587	580
Taiwan	569	570	551	611	614	609
South Korea	554	570	551	616	617	612
Japan	541	561	568	558	574	579
Thailand	443	451	453	423	428	429
Malaysia	403	424	439	444	439	426
Indonesia	402	398	413	378	384	388
Results obtained from (Martin et al., 2012; Mullis et al., 2012)						
			2015			
		Science		**Mathematics**		
Country	**Knowing**	**Applying**	**Reasoning**	**Knowing**	**Applying**	**Reasoning**
Singapore	594	600	595	633	619	616
Hong Kong	547	541	550	600	595	591
Taiwan	589	565	560	598	602	602
South Korea	555	552	560	607	606	608
Japan	567	575	570	578	592	591
Thailand	469	450	447	425	431	435
Malaysia	466	476	467	472	463	453
Results obtained from (Martin et al., 2016; Mullis, Martin, Foy, & Hooper, 2016)						
			2019			
		Science			**Mathematics**	
Country	**Knowing**	**Applying**	**Reasoning**	**Knowing**	**Applying**	**Reasoning**
Singapore	621	608	595	614	614	620
Hong Kong	501	501	504	580	575	582
Taiwan	600	567	559	616	610	616
South Korea	558	560	564	614	604	609
Japan	563	576	570	589	596	599
Malaysia	442	473	459	451	464	462
Results obtained from (Mullis et al., 2020)						

In addition to the poor performance of Malaysian students in international assessments, unemployment in Malaysia is another issue with yet more far-reaching implications. The Department of Statistics Malaysia (2011) reported that the unemployment rate among Malaysian graduates rose from 3.2% in 2007 to 3.7% in 2009. The unemployment rate then declined slightly to 3.5% in March of 2016 (Department of Statistics, 2016; Hossain, Yagamaran, Limon, Nasiruzzaman, & Karim, 2018).

From an economic perspective, these unemployment rates suggest that the labour force in Malaysia is not fully utilised (Nazron, Lim, & Nga, 2017; Razak, Ahmad, & De Mello, 2014; Zhao, 2012). Interestingly, researchers investigating job vacancies and placements in Malaysia noticed that job opportunities are growing yearly. Despite this growth, these jobs are occupied by only a small fraction of the workers, suggesting that the high unemployment rates in Malaysia are not caused by a lack of job opportunities. Instead, unemployment is sustained by the low employability skills among the graduates, especially first-degree students in public higher education institutions who are incapable of fulfilling existing job vacancies. Hence, university students are now entering an era where they need to apply for multiple jobs and go through a long and challenging process to gain employment (Abd Majid, Hussin, Norman, & Kasavan, 2020; Fazaniza, 2016; Hanapi & Nordin, 2014; Nooriah & Zakiah, 2017; Seng, 2018).

Several authors have attempted to conduct studies to characterise the state of the field, oftentimes reaching similar conclusions regarding the reasons behind low employability rates. For instance, according to Mustapha and Greenan (2002) and Hanapi and Nordin (2014), graduates in Malaysia are competent in technical skills but deficient in motivational, critical thinking, problem-solving, communication, interpersonal, and entrepreneurship skills. Furthermore, Downe, Loke, Ho, and Taiwo (2012) and Ismail, Yussof, and Sieng (2011) observed that employers want a more flexible workforce with advanced soft skills such as critical thinking, problem-solving, and analytical thinking in order to perform adequately in the workplace and compete in the global arena. Meanwhile, Ngoo, Tiong, and Pok (2015) found that employers desire their workers to be independent and leaders rather than followers. They also require their workers to have various soft skills, such as critical thinking and problem-solving capabilities, professional and moral skills, and leadership competencies. The authors called on universities to develop curriculums based on the market's demands to help students nurture and acquire these skills under the mentorship of well-informed academics. In addition, they urged graduates to adapt themselves to meeting market demands concerning the development of these skills.

2.6 Barriers to Critical Thinking

Previous research has outlined several obstacles that impede students' capacities to develop critical thinking capacities. For the purposes of this study, the most common barriers are explained below.

According to numerous studies, the memorisation of knowledge is highly detrimental to the development of critical thinking. For example, Ozkan-Akan (2003) conducted a survey that consisted of 522 teachers in four different regions of Turkey. This survey aimed to examine teachers' perspectives on the constraints that affect the development of students' critical thinking in schools. The findings revealed that 73% of teachers (25.5% strongly agreed and 47.5% agreed) noted that the curriculum is designed in a way that leads solely to memorisation of knowledge. Moreover, 76.2% (33.7% strongly agreed and 42.5% agreed) of teachers argued that the curriculum is not designed to enhance students' critical thinking.

In a more recent investigation in the United States, Reynolds (2016) conducted a study comprising 54 junior and senior high school teachers. Reynolds attempted to identify the issues that teachers believed could hinder the integration and utilisation of critical thinking in schools. Feedback from teachers showed that 63% agreed with the claim that the curriculum encourages students to memorise information and facts. The study also showed that 57% of teachers indicated that the curriculum does not address the importance of developing and advancing students' critical thinking capabilities.

The second barrier is ‘teaching to the test’. Standardised testing began as a bid for educational equity and, today, involves the same test being administered in the same way to all students. Standardised testing is implemented in such a manner that the questions, conditions for conducting, marking methods, and interpretations are consistent and predefined (Graf-Webster, 2011; Shelton & Brooks, 2019). However, having all students participate in the same standardised test is similar to asserting that we have ‘standardised’ children whom we expect to acquire knowledge and experience in the same ways and demonstrate their learning experience in the same fashion: through these standardised examinations (Dotson & Foley, 2017).

The ‘teaching to the test’ method has received significant criticism for several reasons. For instance, Dewey (1938) remarked that learning information solely to pass a test will not prepare a student for the analysis and resolution of problems in life. Other scholars have perceived this educational practice as a method that prevents the incorporation of critical thinking skills into classroom environments (Crenshaw, Hale, & Harper, 2011; Kanbay & Okanlı, 2017; Landsman & Gorski, 2007; Sandholtz, Ogawa, & Scribner, 2004; Snyder & Snyder, 2008; Wong, 2007). The ubiquitous utilisation of standardised testing relates to the reliance on examination results as the only variable for making impactful decisions, such as whether to keep students in a specific grade, allow them to graduate from high school or university, and reward teachers and schools with finance, rankings, or ratings (Magee & Jones, 2012). The U.S. Department of Education (2015) addressed these concerns by recognising that it had participated in the present culture of excessive testing and the unnecessary stress resulting from high-stakes testing; in response, the department suggested ‘fewer and smarter’ examinations.

Another major barrier to critical thinking is the limitation of time. In Ozkan-Akan’s (2003) study, 70.7% of teachers cited insufficient time as a reason for not using critical thinking in classrooms. Kowalczyk, Hackworth, and Case-Smith (2012) addressed the importance of critical thinking capabilities within the school environment and teachers' confidence in applying those skills. The researchers learned that insufficient time for instructors to learn new teaching methods was one of the main obstacles to implementing critical thinking skills within the classroom. Furthermore, a recent study by Wegrzecka-Kowalewski (2018) found that, although instructors understand the significance of providing students with the means to actively learn on their own, they are under enormous pressure to finish curricular activities on time and teach students 'to the test'. Several teachers acknowledged that they do not teach students how to think independently, ask questions, and think through solutions.

2.6.1 Barriers to Critical Thinking in Malaysia

In Malaysia, teachers and educators face the same barriers to integrating critical thinking in education. For example, Said, Mohsin, and Yunus (2008) identified the educational system as the main reason teachers and instructors cannot introduce and embed critical and creative thinking abilities or Kemahiran Berfikir Secara Kritis dan Kreatif in their schooling practices. The researchers argued that the current examination-oriented system is why teachers do not view student-centred learning methods favourably. Fieeq (2011) observed that the examination-oriented style of education had been transformed into a phenomenon called 'Kejar Syllabus', which translates to 'chasing the syllabus'. Students often experience this phenomenon when their teachers rush to cover the prescribed learning material before examinations.

To investigate this issue further, several researchers from the Higher Education Leadership Academy or Akademi Kepimpinan Pengajian Tinggi at the Ministry of Higher Education examined 125 lessons from 41 schools across Malaysia. Their investigation found that 12% of classes were delivered at a high standard and incorporated effective pedagogies, and 38% of lessons met satisfactory standards. However, 50% of the lessons were given in an unsatisfactory manner. The researchers noted that teachers placed significant attention on ‘surface-level content understanding for summative assessment purposes, rather than on cultivating higher-order thinking skills’ (Kementerian Pendidikan Malaysia, 2013, pp. 5–2).

With respect to other barriers, such as training and experience in critical thinking, Said et al. (2008) noted that some teachers shy away from teaching critical thinking as they think such teaching approaches will affect their ability to control the class. This issue also caught the attention of local media: on 25 July 2010, *The Star Online* reported that teachers in Malaysia still do not have the time to introduce creativity in their classrooms, and the educational curriculum does not provide teachers with the freedom to adopt critical thinking (Zainal Shah, 2011). A study by Choy and Cheah (2009) noted that many instructors in Malaysian higher education institutes lack a sufficient understanding of critical thinking and the expertise to incorporate it into their teaching practices. According to the instructors, the key reasons behind these issues are the structures of the traditional method of teaching and the lack of training to use innovative teaching techniques. As a consequence of the current practices, the majority of learners in Malaysia are not taught how to search for and utilise information: thus, they become too reliant on their teachers in what is known as ‘the spoon-feeding phenomena’ (Mohtar, Halim, Samsudin, & Ismail, 2019; Zainal Shah, 2011).

In another study, Yusuf and Shah (2018) endeavoured to place more emphasis on the barriers that affect the integration of critical thinking in Malaysian classrooms. Responses from 40 Malaysian primary school teachers revealed that teachers recognise the importance of critical thinking and the role of such a skill in the 21st century. Nevertheless, the researchers identified two types of barriers that exist in the Malaysian classroom. The first was defined as teacher-related barriers. The authors learned that teachers' concerns were primarily centred around the following obstacles: (a) teachers feel a need to cover content, (b) teachers do not have enough time to prepare for developing activities towards critical thinking skills, and (c) teachers do not provide sufficient time for critical thinking in class. The second type refers to student-related barriers. According to Yusuf and Shah (2018), such barriers hindering the adoption of critical thinking practice include the fact that (a) students lack experience in improving critical thinking skills in school, (b) students expect that each question has one correct answer only, (c) students are afraid of making mistakes, and (d) students lack the necessary background knowledge to improve their critical thinking skills.

2.7 Summary of Critical Thinking Skills

Based on the information presented in previous sections, promoting students' critical thinking has the following advantages.

- Enhancing critical thinking allows students to gather information from multiple sources, analyse the authenticity of the collected data, synthesise the results, and make a sound and just decision.
- Improving Malaysian students' critical thinking competencies will significantly enhance their job prospects as most employers are looking for graduates capable of facing the increasing demands of the 21st-century market, which requires the ability to work collaboratively and competitively and think outside the box.
- Effective critical thinkers will be able to achieve the government vision aiming to transform Malaysia into a developed nation.

Additionally, the traditional method of teaching has been identified as a core element that impedes the advancement of critical thinking among students because it has the following drawbacks.

- It is a teaching practice that encourages or fosters rote memorisation.
- It is a teaching practice that has barely evolved over the past few decades.
- It places a significant emphasis on assignments and tests.
- Students are mainly evaluated based on their test scores, which places unprecedented pressure and stress on students.
- From the teachers' perspective, the traditional method is a time-consuming process that strongly affects educators' abilities to teach students how to think critically and creatively.
- Locally, Findings from the PISA and TIMSS demonstrated that students in Malaysia lack the ability in doing problem-solving. By reviewing these findings over the years in multiple domains and reflecting on the teaching and learning practices, it can be concluded that the students lack the opportunity and exposure to develop higher-order thinking skills.

2.8 Effective Teaching Methods

The 2015 Gallup–Purdue Index report found that students were one and a half times more likely to feel that their education was worth the cost if they experienced an opportunity to apply what they had learned in the classroom (Daniels, 2015). Unfortunately, yet unsurprisingly, the traditional methods may contribute to the weak transition from the classroom to professional life because, in these methods, the focus is primarily on what students should learn, instead of how students can obtain skills and knowledge to apply, in turn, to practical problems (Agdas, 2013; Shepherd & Cosgrif, 1998; Sumirattana, Makanong, & Thipkong, 2017).

Many empirical and theoretical studies have been conducted to address the limitations of the traditional teaching method and identify other approaches that could promote students' critical thinking. As a result, several active and technology-enhanced learning methods are available and have been used, to some extent, to encourage students' critical thinking. Researchers believe that when using such methods, an intellectual student–computer partnership is formed, whereby the utilisation of computers extends and amplifies the student's thinking capacity. This partnership makes learners think deeply about the issue at hand, which leads to the generation of thoughts and solutions that would be difficult to achieve without the existence of this partnership (Leen, Hong, Kwan, & Ying, 2014).

In the present study, the researcher focuses on one of the most encouraging technology-enhanced and active learning methods, DGBL, which has been found to be promising and effective in not only improving students' knowledge acquisition, motivation, and attitudes towards learning but also equipping them with much-needed skills, such as critical thinking. DGBL motivates students to take a more productive approach towards studying and learning.

2.9 DGBL Overview

In a survey of 80,000 students conducted by researchers at the University of Indiana in the United States, two-thirds of the surveyed population reported that they feel bored in school at least every day (Gillispie, Martin, & Parker, 2009). Consequently, researchers began to explore other learning methods that engage and interest students. One of the teaching methods most capable of motivating students to learn is via games. Games have been utilised for learning purposes for centuries (McGonigal, 2011; Wilkinson, 2016) and have evolved dramatically from basic dice games in Ancient Greece to sophisticated high-tech computer games (Palmer, 2016). The role that games play in the learning process was even recognised by Plato, who advocated ‘learning math through games’ (Alexander, Eaton, & Egan, 2010, p. 1835).

With the advent of digital computers, rapid development in ICT technologies, and the increased popularity of video games among children and young adults, several scholars have proposed that the next revolution in teaching, learning, and training will be based on DGBL (Barab et al., 2005; Gee, 2003; Prensky, 2007; Xu et al., 2019). The effectiveness of DGBL is attributed to its versatility and ability to motivate a young generation used to playing video games (Prensky, 2003). Furthermore, games can make a dull subject more dynamic by immersing students within an enjoyable environment, thereby motivating them to learn (Behnamnia et al., 2020; Breien & Wasson, 2021; Jackson & McNamara, 2013).

According to Menn (1993), students remember about 10% of what they read, 20% of what they hear, and 30% of what they see. However, if they watch someone model something while explaining it, then their recall jumps to 50% and almost 90% if they are able to participate in the job themselves, even if it was through a simulation. Therefore, the nature of learning in DGBL is different from that of the traditional method. In DGBL, the student is at the heart of the learning process, whereas in the traditional form, students are mostly passive and act as mere recipients of information (Hackathorn et al., 2015; Michel et al., 2009). Moreover, gaming helps students understand concepts much more deeply and often spawns longer and more content-rich answers. This is different from a lecture style or memorisation task in which students would normally answer simple multiple-choice questions (Lang, 2014). Finally, in DGBL, when students interact with the application, their learning becomes stealthy: that is, students do not realise they are learning embedded content, as such, increasing their knowledge and comprehension in the subject being taught (Annetta, 2010; Chang & Hwang, 2019). Table 2.4 highlights the main differences between the DGBL approach and the traditional method of teaching.

Table 2.6: Differences between DGBL and Traditional Approaches

DGBL	Traditional teaching
DGBL allows students to participate in a learning activity (Hwa, 2018).	Students are passive learners with little to no role in the learning process (Liu & Long, 2014).
It helps students retain information better than the traditional method (Wouters, Van Nimwegen, Van Oostendorp, & Van der Spek, 2013).	Traditional teaching is less effective than DGBL in terms of information retention (Wouters et al., 2013).
Students' needs and demands are satisfied and increased through playing (Hwa, 2018).	Information is delivered uniformly, and students' needs are rarely taken into consideration (Liu & Long, 2014).
DGBL keeps students immersed and engaged (Chu & Chang, 2014) and encourages creativity and self-expression (Hwa, 2018).	Students often lose interest during the first 15–20 minutes of the lessons. Traditional teaching encourages rote learning and memorisation of facts (Reynolds, 2016).

Despite these core differences, scholars have proposed parallels between DGBL and traditional teaching approaches. For instance, Apperley and Walsh (2012) and Salute (2015) argued that obtaining information and knowledge from a DGBL application could be more effective than gaining the same information by reading due to the more active role students assume when they use a DGBL application. They also noted, however, that when reading books, different readers can interpret the text in unique ways; they can imagine and reconstruct the book's events and form a distinct understanding of the plotlines. Similarly, when playing a game, individual learners can have completely different experiences and interpretations out of the game's storyline.

2.10 DGBL Definitions

The origins of 'serious games' can be traced to a book of the same name written by Clark C. Abt and published in 1970 (Djaouti, Alvarez, Jessel, & Rampnoux, 2011; Wilkinson, 2016). In his book, Abt defined serious games as "games [that] have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement" (Abt, 1971, p. 9). Although Abt is known for coining the term 'serious games', it was Sawyer's white paper *Serious Games: Improving Public Policy Through Game-Based Learning and Simulation* in 2002 that brought the term to a bigger audience and wider circulation (Djaouti et al., 2011; Sawyer & Rejeski, 2002; Susi, Johannesson, & Backlund, 2007). In the following years, Sawyer recalibrated the definition of serious games to "any meaningful use of computerised game/game industry resources whose chief mission is not entertainment" (Sawyer, 2007, p. 12).

The definitions cited in this section are a mere selection from a much larger number of definitions. While these classifications differ from one another in some areas, most of them share one major component: that serious games are gaming applications used primarily for more than leisure or a time-passing activity. For example, Prensky's (2001) definition of serious games was "Entertainment Games with Non-Entertainment Goals"; a more detailed definition was provided by Zyda (2005, p. 26), who viewed serious games as something "more than just story, art, and software, however. . . . They involve pedagogy: activities that educate or instruct, thereby imparting knowledge or skill. This addition makes games serious". Michael and Chen (2005, p. 21) stated that serious games are "Games that do not have entertainment, enjoyment or fun as their primary purpose". Finally, Mayer and Johnson (2010, p. 245) defined educational computer games as a medium "that is intended to cause a desirable change in the player's knowledge".

The term 'serious games' is synonymous with DGBL. Thus, most research studies – including the present study – will use these two terms interchangeably (Corti, 2006; Squire, 2008).

2.11 DGBL Research

Despite the abundance of research on DGBL, many teachers and instructors still debate the educational value of serious games and whether DGBL can improve students' learning, attitude, and motivation (Hovious, 2015). DGBL has been undervalued by science educators, administrators, policy makers, and parents (An, 2018; Ecker, Müller, & Zylka, 2011). As a consequence, scholars have conducted numerous reviews and meta-analyses to address this concern and provide empirical evidence pertaining to the instructional benefits of this teaching method.

2.11.1 Research on Science in DGBL

Governments across the world in developed and developing nations constantly emphasised the significance of developing and advancing students' science competencies and literacies due to its essential role in achieving people's well-being and assisting them to fulfil their economic, environmental, and social goals (Falloon, 2017).

However, prior research works revealed that students' interest in the science subject is declining (Badri et al., 2016; Iqbal, Bibi, & Iqbal, 2015; Samara, 2015). A number of investigations were conducted to underpin the causes of this phenomenon, evidence from these studies suggest that students lack the motivation to learn science (Sadera, Torres, & Rogayan, 2020), perceive the science curriculum to be content heavy and repetitive (Dunlop, Clarke, & McKelvey-Martin, 2019; Osborne & Collins, 2001), and learning activities to be decontextualised (Honey & Hilton, 2011; Mojumder & Keast, 2018). Additionally, students voice their desire to partake in more practical and hands-on learning situations (Cleaves, 2005; Fuad, Deb, Etim, & Gloster, 2018; Owen, Dickson, Stanisstreet, & Boyes, 2008).

According to numerous empirical analyses and case studies, researchers and educators stated that virtual environments employed in DGBL applications could direct and guide students to feel more absorbed and engaged in the learning process (Cheng et al., 2015). Honey and Hilton (2011) declared that gaming applications have an encouraging capability to promote critical attributes associated with science learning. Further, Cheng, Su, Huang, and Chen (2014) and Zydney and Warner (2016) argued that some of science learning is better studied outside the classroom, in its natural environment. Such features of science learning can be freely experienced by students in interactive microworlds and gaming adventures. Other aspects of science learning are not visible with the naked eye and require graphical representation and visualisation for students to ensure a more appropriate comprehension of the issue being studied.

2.11.2 Research on DGBL Applications' Effects on Students' Achievement

This section of the literature review examines a number of systematic reviews and meta-analyses of research on the instructional value of gaming applications for students' achievements and engagement. For instance, Vogel et al. (2006) carried out a meta-analysis in which the authors investigated the efficacy of games and interactive simulations in the classroom compared to the traditional method of teaching. To achieve this goal, Vogel and her colleagues reviewed research from 1986 to 2003 and included 32 articles for review out of the 248 studies evaluated during the initial screening process.

The outcomes of their meta-analysis indicated that, regardless of their educational level, students who used gaming applications or interactive simulations achieved higher cognitive gains than their fellow classmates who were taught via the traditional teaching method. With regard to learning dynamics (e.g., individual and collaborative), Vogel et al.'s analyses suggested that both learning strategies were superior to those of the traditional method; nonetheless, students learning individually rather than in teams were more likely to achieve better cognitive gains. The researchers reported that these outcomes would be useful for scholars and educators interested in the effects of gaming on students' learning. However, they cautioned that further studies are needed to verify the reliability of their results.

Connolly et al. (2012) also performed an extensive review to examine the impact of DGBL applications on students' achievements and motivation. The authors utilised a multi-dimensional framework that categorised the outcomes of serious and entertainment games into the following groups: knowledge acquisition, affective and motivational skills, perceptual and cognitive skills, motor skills, soft skills, social skills, behaviour change, and physiological.

Their analysis, which was derived from 192 papers published between 2004 and 2009, suggests that there is reason to be optimistic about gaming in education. Further investigation demonstrated that different gaming genres (e.g., role-playing, adventure, board games, etc.) were used to teach students a wide range of curricular subjects with a central emphasis on promoting students' knowledge acquisition. The authors stated that although the utilisation of the multi-dimensional framework helped them to explore various aspects of gaming in education, there remains a need to conduct more randomised controlled trials (RCT) studies comparing the effectiveness of DGBL applications with the traditional method to provide additional evidence regarding DGBL's efficacy in enhancing students' learning gains and motivation.

Moreover, Young et al. (2012) implemented a systematic literature review to study the influence of playing games on learning in classroom environments. The researchers reviewed 39 articles out of more than 300 videogaming studies and classified the works into five sections: mathematics, science, languages, physical education, and history. Their findings showed that gaming was an effective strategy in certain subjects, such as languages and physical education, but there was a notable dearth of empirical evidence concerning the influence of DGBL applications on students' learning achievement, particularly in science and mathematics. Young et al. (2012) argued that the limited effectiveness of gaming applications in science and mathematics could be attributed to a disconnect between the efficacy of the games and their usefulness in the classroom environment. In addition, they noted that many educational video games lack clearly defined learning objectives and outcomes. Consequently, they concluded that there is a need for more sophisticated empirical studies into science-specific instructional video games to understand their impact on student learning and achievement.

Abdul Jabbar and Felicia (2015) performed a systematic literature review and analysed 91 research studies published between 2003 and 2013. The authors aimed to address the limited research on the influence of game design on learning outcomes and identify how the design of DGBL applications could drive students' engagement and learning. The result of their synthesis indicated that, in most cases, research shows that gaming not only provides learners with something to attain from the process of gaming but also enables students to broaden their knowledge and enhance their ability to exercise the learning experiences offered by the application. Abdul Jabbar and Felicia (2015) further articulated that by researching the gaming features that cause enjoyment and motivation, researchers will be capable of recognising what makes students immersed and withdrawn from gameplay and learning. This can be observed by students being confident and driven learners who wish to access and comprehend the content of the gaming application to fulfil its tasks and achieve their goals. Hence, these gaming applications should avoid causing frustration in students, which might hinder their learning or negatively affect their attitude.

Boyle et al. (2016) conducted a systematic examination of research works on the instructional effectiveness of gaming applications in education in a bid to update the findings of Connolly et al. (2012). A total of 512 studies were reviewed at the initial screening process; however, only 143 satisfied the final inclusion conditions and were incorporated in the review. The authors utilised a multi-dimensional framework, and their findings supported those of Connolly et al. (2012) concerning the effectiveness of serious games in promoting students' knowledge in different subjects. These findings suggest that there is an increasing interest in researching the effects of gaming applications in education. Nonetheless, the similarities in Boyle et al.'s (2016) and Connolly et al.'s (2012) results imply that researchers are reusing the same research questions and using DGBL applications for the same purpose.

Furthermore, Clark et al. (2016) reviewed 68 research studies published between 2000 and 2013. The works focused on comparisons of serious games versus non-game conditions (e.g., the traditional method) and augmented (i.e., a gaming application equipped with a special learning or design feature) versus standard games (i.e., the same gaming application but without the special feature). The results from these comparisons indicated that DGBL applications significantly enhanced students' cognitive, interpersonal, and intrapersonal learning outcomes when compared to the outcomes of students in non-gaming conditions. Additionally, Clark et al. (2016) found that DGBL applications with augmented learning designs were linked with higher learning results than DGBL applications without the added design features (i.e., standard serious games). These findings reinforced the notion that DGBL applications are better suited to improving students' learning performance than the traditional method. Meanwhile, the outcomes of this meta-analysis highlighted the significance of instructional design in promoting students' learning and engagement.

As noted previously, in their extensive reviews, Boyle et al. (2016) and Connolly et al. (2012) focused solely on the effects of gaming applications on the learning achievements, engagement, and motivation of post-elementary students. To address this paucity of research, Hainey et al. (2016) reviewed 105 research studies published between 2000 and 2013 that aimed to evaluate the instructional effects of DGBL applications in elementary education. The authors of this review, who also used the multi-dimensional framework, stated that in elementary education, different gaming genres were utilised to instruct several curricular subjects and that most games were used to improve students' content understanding. Their outcomes also suggested that gaming showed a considerable promise in elementary education. However, they stated that more rigorous research is needed with a particular focus on studies that follow the RCT approach and more comparisons between gaming applications and the traditional teaching method. In addition, more research works are needed to ascertain the effectiveness of DGBL applications in collaborative learning conditions.

The findings drawn from these systematic literature reviews and meta-analyses add to the growing evidence that serious games are effective learning tools, and their use is linked to higher learning gains, improved motivation, and learning attitude, as shown in Table 2.5. However, these findings should be cautiously interpreted for several reasons. Firstly, these studies focused on several curricular subjects, such as mathematics, science, and languages, and spanned multiple educational levels without addressing the specific instructional needs and requirements of a single subject area or a specific educational level, thereby making it difficult for researchers and educators to determine the impact of DGBL on a given subject or educational level. Secondly, a considerable number of the studies included in these reviews were either qualitative studies that presented their findings in an anecdotal fashion or case studies that lacked a control group. Finally, as mentioned earlier, although most studies reported positive findings, a closer analysis of these findings implies that research in the domain of DGBL has focused primarily on using serious games to facilitate students' knowledge construction and promote their understanding of curricular information, which left other areas of significant pedagogical interest, such as students' 21st-century skills, largely unexplored.

Table 2.7: Summary of Major Systematic Reviews and Meta-Analyses

Author	Timeline	Number of papers	Target audience	Conclusion
Abdul Jabbar & Felicia (2015)	2003–2013	91	8–14 years old	DGBL allows students to learn in a self-centred learning environment.
Boyle et al. (2016)	2009–2014	143	Post-elementary (over the age of 14 years)	DGBL promotes students' learning comprehension.
Clark et al. (2016)	2000–2012	68	K–16 (aged 6–25)	Digital game interventions were more effective than non-game conditions.
Connolly et al. (2012)	2004–2009	129	Post-elementary (aged 14+)	DGBL showed positive effects on learning construction and motivation.
Hainey et al. (2016)	2000–2013	105	Primary education	DGBL provides students with a promising and effective learning approach.
Vogel et al. (2006)	1986–2003	32	Preschool to higher education	Computer games and interactive simulations resulted in better cognitive gains.
Young et al. (2012)	N/A	39	K–12 (elementary to higher education)	DGBL constitutes an effective learning tool in some areas of education, such as language learning; limited evidence to support the effectiveness of DGBL application in the domain of science.

2.11.3 Research on DGBL Applications' Effects on Students' Science Achievements

Although the previous literature reviews and meta-analyses provided valuable insights into the effects of DGBL on learning, DGBL's influence on specific subjects was inconclusive. Therefore, several scholars have looked closely and critically at the use of DGBL applications in the teaching of science.

For instance, Li and Tsai (2013) conducted a review that aimed to examine and assess the effects of DGBL applications in the domain of science learning. A total of 172 articles were included in the first screening process for potential inclusion, but 133 did not employ DGBL applications. As such, only 31 studies published between 2000 and 2011 met the final inclusion criteria. The study generated three significant findings. Firstly, DGBL in the domain of science education has the potential to improve students' learning across different educational levels. However, the majority of interventions were conducted at the post-elementary and post-secondary levels, and only a few studies were implemented at the primary level. Secondly, most research works have emphasised implementing DGBL applications to facilitate the learning of curricular-based scientific knowledge and scientific concepts. Finally, the research in this review falls under two categories. The first is case studies, in which the research is either evaluated qualitatively or quantitatively without the presence of a comparison group; thus, the research featured one group only. The second category followed a pairwise comparison approach whereby the outcomes of one group were compared against another to empirically ascertain if DGBL applications were more effective than the traditional method of teaching or other technology-oriented learning methods.

Li and Tsai (2013) also observed that researchers paid little attention to promoting students' problem-solving skills. In fact, they argued that among the articles reviewed, less than one-third targeted students' problem-solving skills. Meanwhile, other skills, such as critical thinking, attracted no research interest. The authors called on researchers to design DGBL applications that correspond to specific learning principles and strategies and then examine the academic effectiveness of those designs.

Cheng, Chen, Chu, and Chen (2015) conducted a systematic review examining the effectiveness of DGBL applications in various science-related areas. The authors retrieved 53 research articles published in 2002–2013, indexed by the Web of Science's Social Sciences Citation Index and Scopus. Their findings indicated that DGBL applications in science education covered several educational levels, with a specific interest in the elementary (17 studies) and junior high school (16 studies) levels, while other educational levels (e.g., kindergarten, senior high school, and university) received considerably less research interest. The synthesis also revealed that the effectiveness of DGBL applications in science was examined by employing quantitative (31 studies), qualitative (eight studies), and mixed-method approaches (14 studies), with the quantitative mechanism being the most commonly used research methodology. Further analysis demonstrated that 43 of the 53 studies included in this review were primarily used to facilitate students' knowledge construction, while 28 of the 53 reviewed articles did not utilise any instructional strategy when the games were designed.

The concluding remarks of Cheng et al. (2015) show that research interest in DGBL applications in the context of science education is increasing. Nevertheless, researchers have to be cautious during the research process as DGBL applications are beneficial only if utilised appropriately. Additionally, the authors of this review advised researchers to consider more studies that emphasise 21st-century skills, such as complex problem-solving abilities, critical thinking, and communication skills, to better determine the overall effect of using DGBL applications in the field of science education.

The synthesis of research presented by Cheng et al. (2015) and Li and Tsai (2013) suggests that DGBL applications can be considered an effective learning strategy in the context of science education. However, a closer analysis shows that these studies covered multiple educational levels and included numerous case studies and qualitative analyses. Therefore, it is difficult to ascertain the extent to which DGBL applications are effective at the elementary level, specifically.

The present study conducted a systematic literature review to confront these concerns and address the following research question: 'What are the potential benefits of using DGBL in elementary science education?' The researcher utilised the Scopus and Web of Science repositories to search for DGBL interventions in science education between 2006 and 2017. These databases were targeted because they are known to index rigorously reviewed research and high-impact research articles. As a result, books and book chapters, conference proceedings, and unpublished dissertations were not considered for inclusion.

The 23 empirical studies included in the review yielded several findings. For instance, DGBL interventions in elementary education have mostly followed a quasi-experimental design. The empirical data also demonstrated that serious games were used for different science-related topics, such as plants and insects.

In addition, the research findings suggested that DGBL applications were more effective than the traditional teaching method and that the majority of the applications were used to improve students' knowledge construction and information retention. Previous general and domain-specific review findings corroborate these observations by indicating that DGBL applications could promote students' learning and motivation.

Furthermore, the review generated several recommendations. Firstly, additional RCT research works are needed to examine the impact of different learning dynamics (e.g., individual and collaborative) on science learning. Secondly, future research should assess how serious games could affect students' learning and performance in other domains, such as critical thinking, creativity, and complex problem-solving abilities.

Based on the information derived from the research reviews in this section, DGBL applications in science learning have three main advantages. Firstly, serious games play a critical role in promoting students' achievements and engagement in science learning. Secondly, to assist young learners in going beyond learning science to actually 'knowing' the science, their instruction should not focus on memorisation of facts; rather, their teaching should focus on a deeper understanding of science. Hence, the affordances native to serious games may provide a learning environment in which learners can develop a better understanding of science because they can actively participate in the experiment and visualise how the variables of this experiment interact. Finally, DGBL applications have provided students with the opportunity to engage with science-based situations and exercises in the safety of a virtual learning environment.

By taking this domain-specific approach, the findings of these reviews also indicated that most research focusing on the science subject emphasised students' knowledge construction. Conversely, other areas, such as critical and creative thinking, received little to no research attention.

2.11.4 Research on the Effect of DGBL Applications on Students' Critical Thinking in Science

The previous two sections reviewed and investigated research works on the impact of DGBL applications on students' achievements, motivation, and attitude. The findings of research in these domains observed a positive effect of DGBL applications in different learning environments. However, studies concerning the impact of this method of learning and training on students' critical thinking and other 21st-century skills, such as creativity, communication, and complex problem-solving abilities, were scarce. As such, a limited number of studies are examined in this section.

One study investigated the effects of playing a serious game on the critical thinking skills of seventh-grade students. Yang and Chang (2013) tried to ascertain if allowing students to design their own DGBL application would help them to improve their concentration and engagement, critical thinking, and academic learning outcomes. Sixty-seven students from biology and programming courses took part in this pre- and post-test quasi-experimental study. Participants were divided into two groups: the experimental group (32; 16 males and 16 females) and the control group (35; 17 males and 18 females). Students in the experimental group were asked to design their own DGBL applications based on biological concepts by making use of the programming experience they acquired in their programming courses, while students in the control group undertook workshops that focused on programming and the development of flash animations based on biological concepts.

By the end of the 19-week intervention, the statistical analyses demonstrated a significant dissimilarity between experimental and control group students in the post-test (p = .01) and delayed test (p = .00). The post-test results of the experimental group's critical thinking assessment (Mean [M] = 19.41, Standard Deviation [SD] = 3.44) was considerably higher than the scores of the students in the control group (M = 17.29, SD = 3.30). Moreover, the experimental group outcome in the delayed test (M = 20.09, SD = 3.56) was better than that of the control group (M = 16.60, SD = 4.67). The findings suggest that using serious games in biology courses improved students' critical thinking skills more than it did their peers in the control group who used flash animations.

The second study compared the effects of two instructional strategies – inquiry-based ubiquitous gaming and conventional inquiry-based ubiquitous learning – on sixth-grade students' learning achievement, motivation, critical thinking tendencies, and problem-solving. In their seven-week long intervention, Hwang and Chen (2017) followed a quasi-experimental research method with pre- and post-test measures to assess the instructional value of the proposed approaches. The 101 (53 males and 48 females) students who participated in this study were divided into two groups. Fifty students were assigned to the experimental group and were taught using the inquiry-based ubiquitous gaming approach, and 51 students were assigned to the control group and instructed using the conventional inquiry-based ubiquitous learning method.

Students' responses to the critical thinking tendency questionnaire indicated a significant dissimilarity between the experimental and control groups ($F = 4.26$, $p < .05$). In the statistical analyses, the experimental group scored a mean of 4.00 and a standard deviation of .65, whereas the control group scored a mean of 3.73 and a standard deviation of .73. These findings showed that students in the experimental group significantly surpassed their peers in the control group, suggesting that students who learned with the proposed inquiry-based ubiquitous gaming approach exhibited considerably higher tendencies to think critically than students in the control group.

Although Hwang and Chen (2017) and Yang and Chang (2013) demonstrated that DGBL could improve students' critical thinking, both studies have a few limitations and drawbacks. For example, in their study, Yang and Chang (2013) utilised the social constructivist theory via the DGBL approach to improving students' critical thinking. This theory is primarily concerned with learners' active construction of knowledge through experiences they have gained and how their interpretations of different contexts will influence their perception of the world (Li & Tsai, 2013; McPhail, 2016). However, according to Knapp (2019), using this theory with young learners could be problematic. This issue may, in part, be linked to students' limited exposure to contextualised learning scenarios that could enrich and improve their social constructivist ideas, especially if they have been taught via the traditional teaching method, where they learn that each question has one correct answer only (Knapp, 2019; Lambert, 2012). Further, Hwang and Chen (2017) attempted to assess students' tendencies to think critically by utilising a self-reported 5-point Likert scale questionnaire, instead of developing a test measure based on the content of students' science curriculum, to examine their actual critical thinking competencies more precisely.

The limited number of studies addressing the relationship between DGBL and critical thinking implies that this research area has received little scholarly attention, particularly in the context of elementary education. In fact, Qian and Clark (2016) conducted a systematic literature review of 29 research studies published between 2010 and 2014 to investigate the effects of DGBL applications on students' 21st-century skills (i.e., the 'four Cs'). Although their findings demonstrated that nearly 70% of the studies included in the review focused on critical thinking, most of these papers referred to higher education, and only one of them assessed the effects of serious games on students' critical thinking capacities in the science subject. This observation was reiterated by Chou, Wu, and Tsai (2019). Their review included 42 research studies published between 2006 and 2017 and focused on the effects of different computer-enhanced learning approaches on students' critical thinking skills. Chou et al. (2019) reported that most studies in this domain focused on higher education (n = 27), followed by high school (n = 9) and graduate students (n = 4), while students at the primary level received the least research attention (n = 2).

This finding is consistent across different learning platforms. For instance, in their attempt to highlight the trends, opportunities, and challenges in mobile DGBL, Chang and Hwang (2019) conducted a systematic review of 113 research studies indexed by the Web of Science online repository and published between 2007 and 2016. The review found that between 2007 and 2016, no study was conducted that focused on primary-level students' higher order thinking skills in the context of mobile DGBL.

2.12 DGBL Frameworks

The main purpose behind using serious games for learning is that such applications are motivational for most people. However, all formal learning should have a foundation in learning-theory and DGBL is not an exception. Evidence from literature suggests that DGBL can be highly effective when used properly in accordance with suitable and effective learning principles (Ott, Popescu, Stanescu, & de Freitas, 2013). In addition, Erhel and Jamet (2013) conducted two experiments that showed DGBL could enhance motivation for learning providing that the design contained "features that prompt learners to actively process the educational content." (p. 156). Further, a number of media-comparison reviews clearly demonstrated that DGBL applications combined with instructional strategies are potentially more effective than DGBL applications designed without such learning principles (Wouters & Van Oostendorp, 2017; Wouters et al., 2013). However, in their extensive reviews, both Connolly et al. (2012) and Young et al. (2012) noted that most DGBL applications are not based on solid theoretical foundations and systematic frameworks that guided their design. This remark is especially true in the teaching of the science subject, where there is a notable dearth of research that aims to investigate the effectiveness of DGBL applications engineered towards specific designs and learning principles (Cheng et al., 2015; Li & Tsai, 2013).

2.12.1 Annetta's Framework

As illustrated in Figure 2.1, Annetta (2010) proposed a framework for serious games comprising six components: identity, immersion, interactivity, increasing complexity, informed teaching, and instructional. The 'identity' aspect deals with the association between the student and the digital character controlled by the learner, which enhances the student's 'immersion' within the gaming scenario. The 'interactivity' component refers to the interaction and communication between the students and other in-game characters or players. 'Increasing complexity' relates to a basic characteristic of gaming applications to sustain students' interest. The 'informed teaching' element refers to the responses and embedded assessments of the DGBL application. Finally, the 'instructional' aspect could be regarded as the aim of the game.

These six components are known collectively as the 'Is framework'. They were chosen after the researcher reviewed numerous articles on DGBL design and development from the primary to the graduate level to provide insights to teachers and instructional designers interested in designing serious games that connect real-world scenarios with textbook content. The six components of this framework were mapped into a high school DGBL application called 'The Great Entomologist Escape'. However, the effectiveness of this game was not empirically evaluated; it is, therefore, difficult to ascertain if the Is framework is a useful or viable serious games framework.

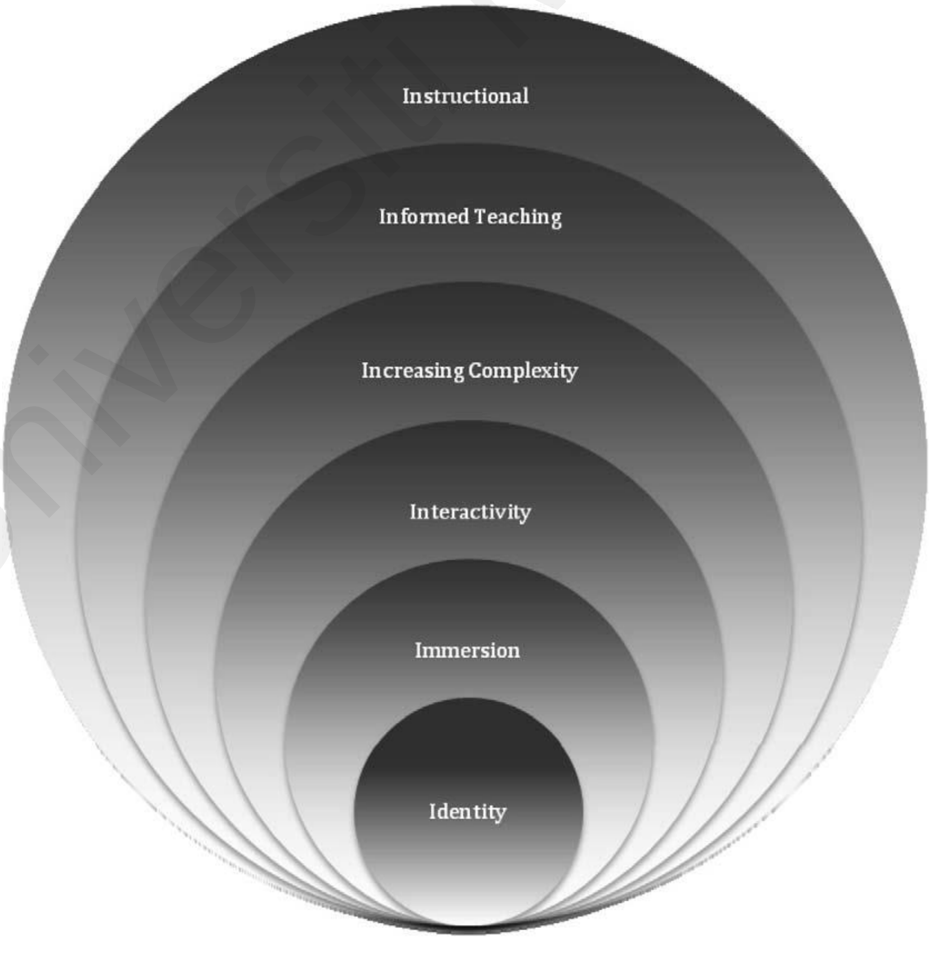

Figure 2.1: Components of the Is Framework (*Annetta, 2010*)

2.12.2 Harteveld, Guimarães, Mayer, and Bidarra's Framework

Harteveld, Guimarães, Mayer, and Bidarra (2010) presented a design framework that aimed to guide the complicated process of designing a DGBL application. As depicted in Figure 2.2., this framework identified three compulsory components for any DGBL application: play, meaning, and reality. The authors argued that these components are equally important, as designers need to create a balance between them to construct a serious game that is realistic, meaningful, and fun to play.

The first component, play, concerns the playability of the game and is mainly associated with the criteria responsible for developing good DGBL applications, including engagement, enjoyment, and immersion; computer technologies and graphics; and specific game elements, such as scores and rules.

The second component, meaning, refers to the game's purposes and is primarily concerned with the learning affordances of relevance, transfer, and reflection. Consideration of these learning affordances ensures that the knowledge gained from game playing is actively processed and comprehended, not merely clicked away to resume playing, as frequently happens in numerous inadequately developed DGBL applications.

Finally, the third element, reality, reflects the degree of correspondence with the real world and its representation within the game. This component is based on the disciplines related to the subject matter of the game: for example, a science game should incorporate elements of biology or physics experiments. This framework guided the development process of a DGBL application called 'Levee Patroller', which was used to teach levee patrollers in the Netherlands. Hence, the game only focused on teaching and educating this limited group of people who are specialised in the adverse effects of water crises.

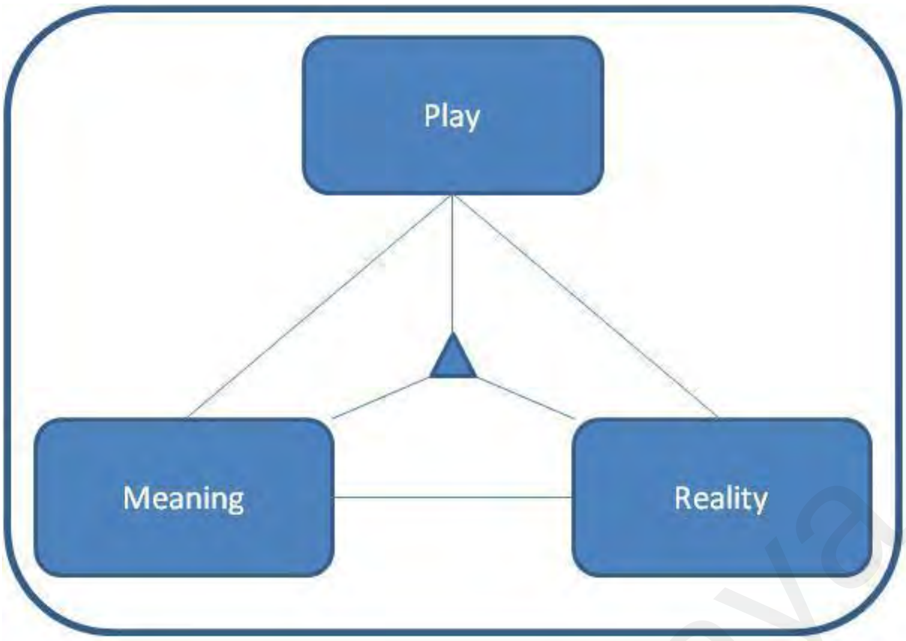

Figure 2.2: The Design Components (*Harteveld et al., 2010*)

2.12.3 Pappa and Pannese's Framework

To promote the sharing of knowledge and transfer for intergenerational learning between younger generations, especially, generation Y and older people, Pappa and Pannese (2010) proposed the e-VITA framework. This framework comprises three components: technical verification, user experience, and pedagogical aspects. This framework perceives the DGBL application as a game, an information technology (IT) product, and a learning instrument. When using this framework to design a DGBL application, designers should ensure that their design is friendly, technically sound, fun, and engaging and serves as an effective learning tool.

2.12.4 Aleven, Myers, Easterday, and Ogan's Framework

Aleven, Myers, Easterday, and Ogan (2010) acknowledged the challenges that lie in developing educational games, irrespective of whether they are digital or not. To address this issue, Aleven et al. (2010) presented a framework aiming to help researchers and game developers to discern and evaluate design-related decisions. The proposed framework consists of three components: learning objectives, instructional principles, and mechanics, dynamics, and aesthetics (MDA). Notably, the latter component is a framework itself. According to Aleven et al. (2010), a game has a higher chance of success if the learning aims are identified in the first stage of the development process. The design also needs to consider game aesthetics and in-game mechanics using appropriate game dynamics (e.g., the pace of the game) and whether the game employs well-grounded instructional design principles. This framework was utilised to analyse a DGBL application called Zombie Division, which is used in primary education.

2.12.5 Linehan, Kirman, Lawson, and Chan's Framework

Linehan, Kirman, Lawson, and Chan (2011) proposed a clear and practical framework for designing engaging educational games: applied behaviour analysis (ABA). The ABA framework comprises three sections, the first of which is made up of two subsections (defining and measuring behaviour). This subsection relates to selecting and clearly defining target behaviours and intended learning outcomes, as well as measuring behaviour, which refers to allocating a numerical value to observed behaviour. The second section of this framework also consists of two subsections: recording and analysing behaviour change. The former entails recording the measurement of learners' behaviour in a way that can benefit both the learner and the instructor. The latter, analysing behaviour change, involves measuring and assessing whether the learner is approaching the desired learning outcome or not. The final section (presenting corrective feedback) is composed of three subsections. The first subsection, presenting feedback, aims to guide the learner to achieve the desired outcomes. The second subsection, evaluating the effectiveness of feedback, concerns observing a student's behaviour and assessing what kind of reward would be the most beneficial for them. The last subsection, scheduling rewards, aims to maintain and maximise students' interest in the learning experience.

2.12.6 Chorianopoulos and Giannakos's Framework

To address the limited number of DGBL frameworks, Chorianopoulos and Giannakos (2014) proposed a serious game framework based on a few established design principles. This framework comprises four components. The first component involves the narrative or story. According to the authors, the rationale of this component is to provide students with the means to practice exemplification and reflection. The second component is popular interactions, incorporated to render games more interesting and familiar to the learner; as such, this element is based on adopting popular game mechanics (e.g., platform games). The third component, trial and error, is grounded on students' abilities to learn from their experiences. Overall, then, this framework offers students immediate feedback during gaming in order to help them identify their mistakes and assess their errors. Moreover, this framework implements collaborative learning among students in an attempt to help them elaborate on and refine the knowledge gained through playing. Finally, Chorianopoulos and Giannakos (2014) used this framework to design two DGBL applications which targetted students in primary education.

2.12.7 Foster and Shah's Framework

Foster and Shah (2015) argued that game-based learning frameworks could help researchers and educators gain insight into students' learning via digital games. The authors employed the inquiry, communication, construction, and expression (ICCE) framework to test this proposal. They stated that educational games need to be designed in accordance with the ICCE guidelines to provide a deeper learning experience. In the ICCE framework, 'inquiry' refers to the knowledge students obtain as they face problems and challenges during their gameplay, while 'communication' refers to students' interaction with their fellow peers or the game itself. The third component of this framework is the construction of knowledge, which occurs when the game helps the learners to demonstrate and express their understanding of a given subject. Finally, the fourth element is expression, which involves recognising the game as a vehicle through which students can express their emotions, feelings, and values. The ICCE framework was implemented in a learning intervention that focused on students in secondary education.

2.12.8 Andreoli et al.'s Framework

Andreoli et al. (2017) proposed and employed a framework entitled FRACH to guide the development process of a virtual reality DGBL application called HippocraticaCivitasGame, which was intended for undergraduate students and adults. FRACH is an iterative framework that extends the ISO 25010 quality standard for software, analysis, and design artefacts by adding two components: 'immersivity' and 'functional collaborativeness'. The ISO 25010 comprises two main sections: quality in use and product quality. As shown in Figure 2.3, 'immersivity' was incorporated into the user satisfaction category under quality in use to increase knowledge acquisition and offer situated learning contexts to learners.

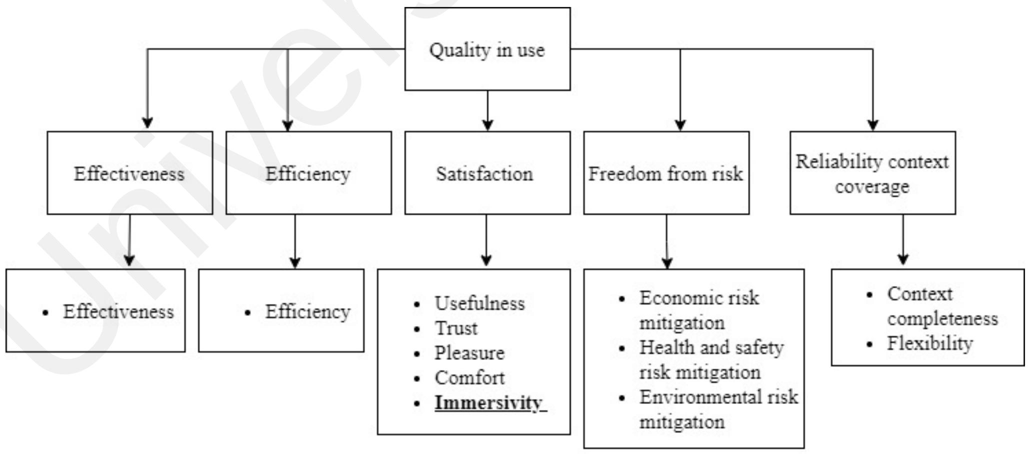

Figure 2.3: The Extension of Users' Satisfaction in ISO 25010 (*Andreoli et al., 2017*)

Functional collaborativeness, meanwhile, was embedded in the functional suitability category under the product quality section, as illustrated in Figure 2.4. This component aims to promote teamwork and information comparisons among learners.

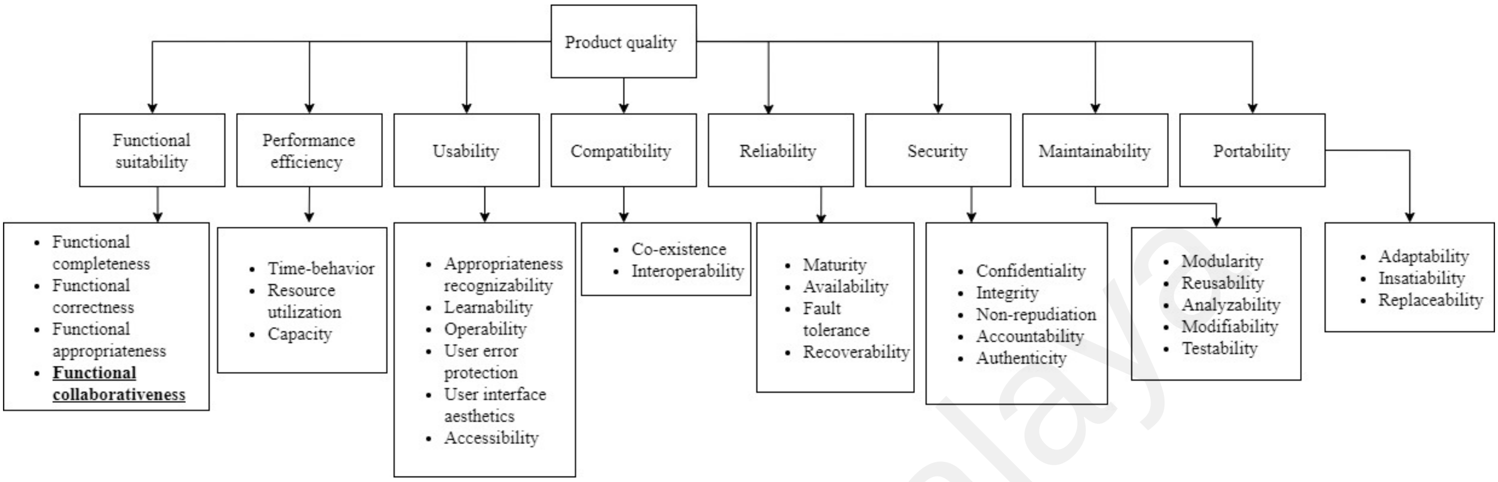

Figure 2.4: The Extension of Functional Suitability in ISO 25010 (*Andreoli et al., 2017*)

2.13 Comparison of DGBL Frameworks

Table 2.6 presents the strengths and weaknesses of DGBL frameworks.

Table 2.8: Comparison Between DGBL Frameworks

Authors	Proposed Framework	Components	Target Users	Strength	Weaknesses
Aleven et al. (2010)	Analysis and design	Learning objectives, MDA framework, and instructional principles	Primary education	Guides the game development process from design and educational angles	It is challenging to use this framework in evaluating and designing educational games due to its highly descriptive nature No empirical evaluation.
Andreoli et al. (2017)	Design framework	Immersivity and functional	Undergraduate students	Validated empirically; emphasises the	Developed primarily for knowledge acquisition;

		collaborativeness		importance of collaboration among learners	the effectiveness of the FRACH framework was evaluated without a control group and by using self-reported questionnaires
Annetta (2010)	Design framework	Identity, immersion, interactivity, increasing complexity, informed teaching, and instructional	Primary to the graduate level	Relinquishes control to the teacher and student, which makes them, in turn, the game's producers	The authors do not show how to integrate these components into the game design; a game was developed, however, no study was found detailing the actual implementation
Harteveld et al. (2010)	Design framework	Play, meaning, and reality	Levee patrollers	Bridges a significant gap in the domain of serious game design	Although it provides a detailed explanation of the issues related to educational game design, it might be difficult for non-experts to follow this philosophy, which could limit its impact The evaluation was carried out via self-reported scales.
Foster and Shah (2015)	Evaluation and design	Inquiry, communication, construction, and expression	Secondary education	Assesses educational games from the perspective of the learners' experience.	The empirical evaluation lacked a control group; developed primarily for knowledge acquisition

				Clear mapping instructions to guide the development process.	
Chorianopoulos and Giannakos (2014)	Design framework	Narrative, popular interactions, trial and error, and collaboration	Primary education	Designed for young learners, especially those at the primary level	The components of this framework lacked detailed descriptions, which could limit developers' abilities to design effective games Additionally, no empirical evaluation of the framework
Linehan et al. (2011)	Design framework	Measuring behaviour, recording and analysing behaviour changes, and presenting corrective feedback	Not available	The structure of this framework resembles that of computer games (e.g., target frequency, which is the continuous practice of a given practice until the goal is met)	Lack of empirical validation
Pappa and Pannese (2010)	Design and evaluation framework	Technical verification, user experience, and pedagogical aspects	Generation Y and older people	The dissection of the design process into three sections allows a designer to recognise and perceive the significance	Lacks emphasis on the game dimension An evaluation was mentioned, however, no outcomes were provided

of these three aspects

Based on the information presented in Sections 2.12.1 through 2.12.8 and Table 2.6, it can be inferred that there are some limitations and gaps that should be addressed to provide a more complementary understanding of the DGBL approach. Some of these limitations are summarised below.

- Most of the frameworks were designed for knowledge construction purposes.
- Most studies lacked empirical analysis to validate the effectiveness of the framework they proposed.
- Most of these studies do not provide procedural guidance in relation to how could other researchers or game developers design DGBL applications based on the frameworks they proposed.

We can, therefore, conclude that there is an evident lack of emphasis on developing DGBL frameworks that investigate the effectiveness of serious games in enhancing students' 21st-century skills, particularly frameworks that focus specifically on critical thinking.

2.14 The Conceptual Process Flow to Derive the DGBL Framework

In the present study, a DGBL framework entitled inquiry, communication, mystery, decision making, challenge, and rewards (ICMDCR) was proposed to address the notable lack of DGBL frameworks that addressed students' critical thinking. Figure 2.5 shows the process involved in deriving the proposed ICMDCR framework.

2.15 The ICMDCR Framework

This section highlights how the ICMDCR framework was proposed to address the challenges of current DGBL applications. This framework was established after conducting an extensive literature review and analysis on critical thinking issues, game design problems, and existing DGBL frameworks. Li and Tsai (2013) observed a dearth of DGBL games that were designed to specifically correspond to a set of learning strategies and principles. Boyle et al. (2016), Cheng et al. (2015), Connolly et al. (2012), and Hainey et al. (2016) also indicated that the majority of the research on DGBL was interested in promoting students' knowledge construction.

The ICCE framework proposed by Foster and Shah (2015) provides a clear mapping and detailed analysis of how ICCE components are incorporated effectively in an educational computer game. However, the ICCE framework could be inadequate for 21st-century educational game design, as it focuses mainly on knowledge construction rather than 21st-century skills such as critical thinking. Therefore, this study proposed the ICMDCR framework, an improved framework that incorporates critical thinking components in educational game design to prepare students for the challenges of modern-day education. The ICMDCR framework avoided the inclusion of the construction component, as it is occasionally linked to reduced motivation to learn and could lead students to experience boredom and disengagement during the learning process (Alyaz, Spaniel-Weise, & Gursoy, 2017; Foster & Shah, 2015).

Moreover, this framework did not include the expression component as its integration can sometimes make students feel frustrated because they cannot modify and configure their character to fit an identity and express aspects of themselves. In this context, students' inability to connect with the game could significantly affect their desire to learn (Birk, Atkins, Bowey, & Mandryk, 2016; Schrader, 2019).

In addition to critical thinking, this framework was designed to address another issue: the balance between entertainment and educational elements. Tahir and Wang (2020) observed that when game designers develop educational games, they emphasise making the game exciting, which could be perceived as them neglecting the educational dimension of the game. Conversely, when educators or teachers develop DGBL applications, they focus primarily on the educational features of the game; as such, they often produce rigid and unappealing games. Therefore, Van Eck (2006) suggested that game developers and educators collaborate. Such a collaboration, according to Van Eck (2006, p. 6), has been perceived as the 'Holy Grail' of DGBL applications development, as this approach recognises the importance of both aspects and can integrate them seamlessly and in equal measures within the gaming application being developed. A detailed description of each of the components is presented below.

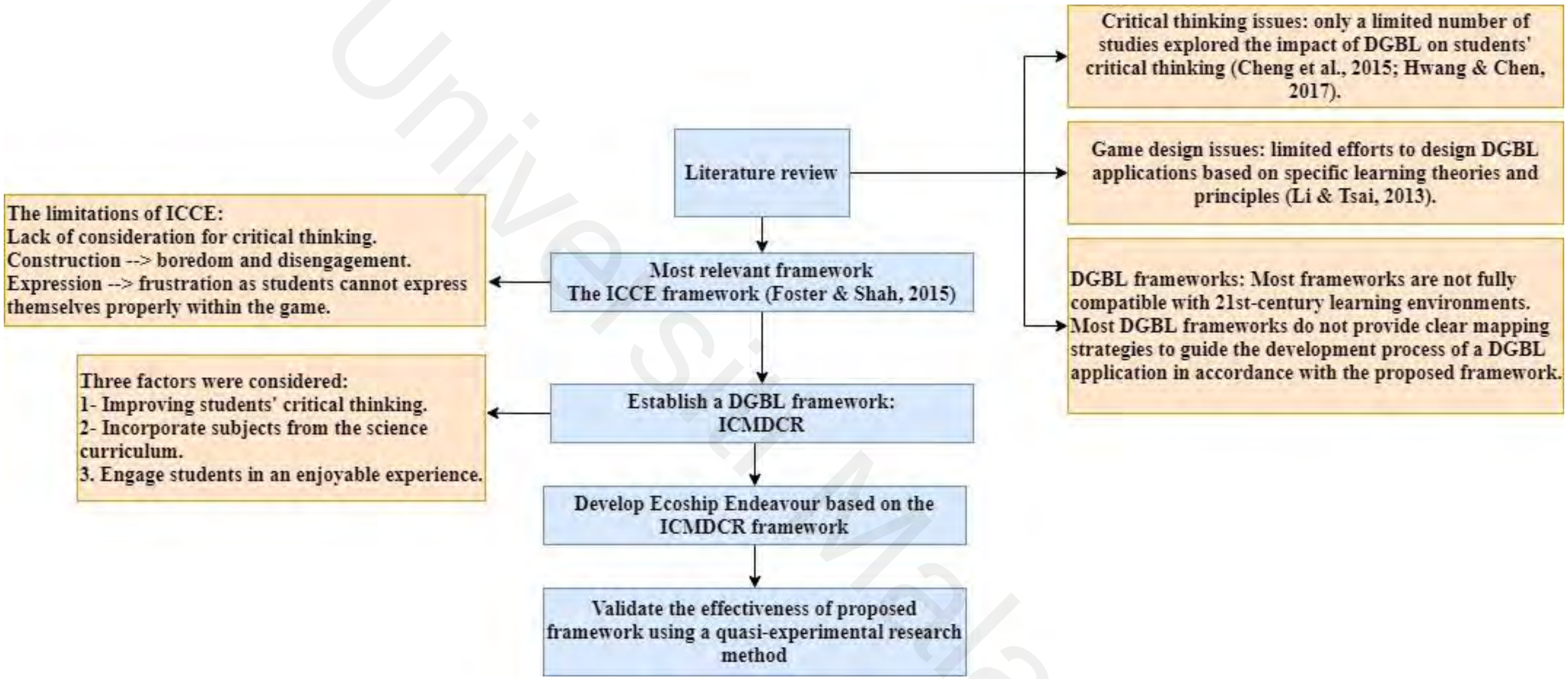

Figure 2.5: ICMDCR Development Process

2.16 Components of the ICMDCR Framework

2.16.1 Inquiry

'Inquiry' is a process in which students acquire knowledge from raising questions and investigating problems (Bruce & Casey, 2012; Chiang, Yang, & Hwang, 2014; Foster & Shah, 2015). Inquiry activities enable students to engage in a self-directed and more meaningful learning environment (Hwang, Chiu, & Chen, 2015; Wang, Duh, Li, Lin, & Tsai, 2014). Researchers have shown that inquiry activity can improve students' understanding of science and scientific practices (Anastopoulou et al., 2012; Donnelly, Linn, & Ludvigsen, 2014), assist in enhancing their critical thinking (Ahmed & Parsons, 2013; Hwang & Chang, 2011; Hwang et al., 2015), encourage the use of higher order thinking and taking responsibility for their learning (Hwang et al., 2015; Oliver, 2008), and promote the transfer and application of knowledge to real-world contexts (Hwang & Chen, 2017; Vogel, Kurti, Milrad, Johansson, & Müller, 2014). Hence, incorporating inquiry-based learning activities into the ICMDCR framework is necessary as it can enhance students' learning achievements, learning motivation, satisfaction degree, and flow state (Hwang et al., 2015), which is a state of complete absorption or engagement in an activity (Csikszentmihalyi, 1991).

2.16.2 Communication

Communication in this framework refers to the interaction between a student and their teammates and with the game. Collaboration with fellow teammates helps to improve students' acquisition of cognitive skills, such as more effective problem-solving, and to raise students' interest in the learning process (Chen et al., 2020; Meluso, Zheng, Spires, & Lester, 2012; Mikropolous & Natsis, 2011). In addition, incorporating this component into the learning process will pave the way for a healthy level of competition among learners, as teammates work together to achieve a common goal and objective (Kennedy, 2014).

In DGBL, interaction with games usually occurs in the form of scaffolds. In an educational game, scaffolds are presented as feedback, hints, clues, backtracking, and explicit instruction (Jabbar & Felicia, 2015; Ke, 2016). Barzilai and Blau (2014) and Yu and Yang (2014) argued that scaffolding is one of the essential features for an effective educational video game design, and it has been found to be appealing to students (Jabbar & Felicia, 2015; Tan, Goh, Ang, & Huan, 2013), reduce the time to solve puzzles, and reduce frustration (Jabbar & Felicia, 2015; Kao, Chiang, & Sun, 2017; Sun, Wang, & Chan, 2011).

2.16.3 Mystery

The role of mystery in educational computer games is crucial and is synonymous with students' curiosity, exploration, uncertainty, and surprise (Alkhafaji, 2018; Sandberg, Maris, & Hoogendoorn, 2014; Shi & Shih, 2015). In an epistemological context, curiosity refers to the level of novel experience a student gains from the learning environment (Billieux et al., 2013; Shi & Shih, 2015). Epistemic curiosity adds to students' satisfaction (Shi & Shih, 2015; Tseng, 2011; Tseng & Teng, 2015). A student's sense of curiosity and exploration is improved when the information presented to them is novel, conflicting, complex, and surprising and when it incapacitates students' future prediction abilities (Sandberg et al., 2014; Woo, 2014). Furthermore, embedding elements of mystery in educational computer games triggers students' motivation and promotes their hypothesis-testing skills (Taub & Azevedo, 2018; Taub, Sawyer, Lester, & Azevedo, 2020).

2.16.4 Decision Making

Decision making is a thinking process that reflects a learner's ability to collect scientific data, make a given number of judgements and evaluations, and select one or more decisions from a larger set of alternatives (Jonassen, 2012; Kim, Anthony, & Blades, 2014; Lin & Lin, 2014; López-Cabrales & Bornay-Barrachina, 2019; Sendurur, 2019). The significance of decision making is evident throughout the literature. For instance, Abdulai and Shafiwu (2014) regarded decision making as one of the skills necessary for a productive work environment. Miri, David, and Uri (2007) stated that students are highly encouraged to improve their higher order thinking skills, which include, among other things, the capacity to make informed decisions. Proficient decision-makers are sought-after employees and are capable of succeeding in the international job market (Partnership for 21st Century Skills, 2011; Yang, Chuang, Li, & Tseng, 2013).

2.16.5 Challenge

Challenge is a significant element that should be considered when designing an educational video game (Chen, 2017; Csikszentmihalyi, 2014; Hou & Li, 2014; Nygren, Sutinen, Blignaut, Laine, & Els, 2012; Hwang, Wu, & Chen, 2012). A challenging educational game that matches the skills of students adds to their feeling of self-efficacy (Chase et al., 2020; Hung, Sun, & Yu, 2015; Power, Lynch, & McGarr, 2020), contributes positively to their learning experience (Chen & Sun, 2016; Liu, Cheng, & Huang, 2011), and improves collaboration amongst learners (Liu et al., 2011; Shih, Shih, Shih, Su, & Chuang, 2010). Challenge can also lead to a better flow experience (Hamari & Koivisto, 2014; Procci, Singer, Levy, & Bowers, 2012; Sung & Hwang, 2013). In addition, a challenging educational video game can positively influence learners' knowledge and concentration skills (Barzilai & Blau, 2014; Hamari et al., 2016; Hou & Li, 2014).

2.16.6 Rewards

When incorporated in a game-based learning environment, rewards can motivate users to learn and can improve their sense of satisfaction (Cicchino, 2015). In a DGBL environment, rewards can take several forms, including currency, rank, mechanical rewards, emotional rewards, and points (Boyle et al., 2016; Bunchball, 2010; King, Delfabbro, & Griffiths, 2011; Simões, Redondo, & Vilas, 2013). The inclusion of rewards in games is an important factor due to the role it plays in engaging learners (Filsecker & Hickey, 2014; Moon, Jang, & Kim, 2011; Westera, 2019). It encourages learners and increases their confidence levels (Hasegawa, Koshino, & Ban, 2015). Moreover, rewards motivate students to explore and complete more missions and carry out multiple attempts to outperform their classmates (Gillispie, Martin, & Parker, 2010; Jabbar & Felicia, 2015; Tan et al., 2013). Figure 2.6 shows the six components of the ICMDCR framework.

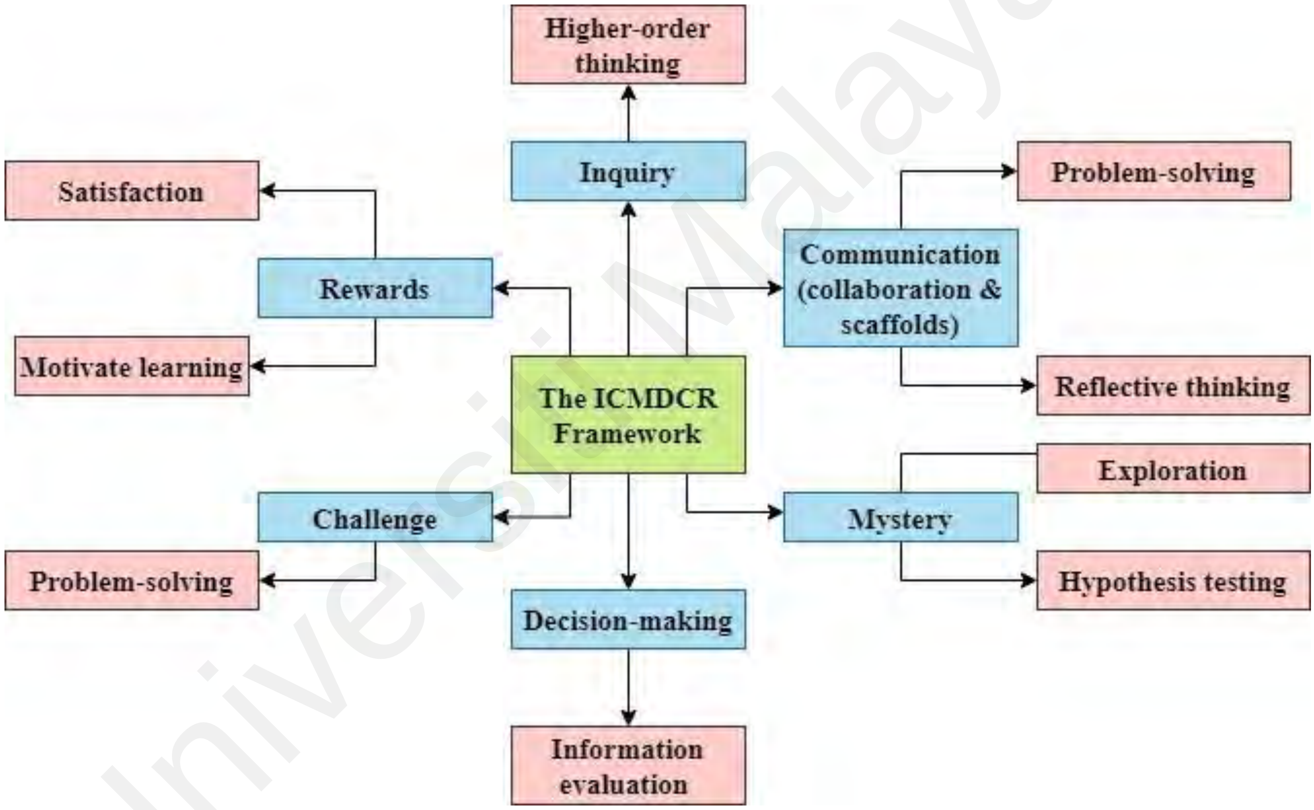

Figure 2.6: The ICMDCR Framework

2.17 Summary

This chapter has shown that the importance of critical thinking has been recognised for more than 2,500 years, and its prominence continues to rise, especially in a 21st-century context. In addition, this chapter highlighted the implications of inadequate critical thinking abilities and explored the barriers that inhibit students' capacity to think critically.

To address these barriers, many researchers have proposed different learning approaches. Among these learning methods, DGBL has been perceived as a promising instructional tool that could promote students' learning and motivation. Many types of studies have been carried out to test the effectiveness of using DGBL applications in the teaching and learning of the science subject. However, most of this research focused on promoting students' knowledge construction and content understanding, and research on 21st-century skills, such as creativity, critical thinking, and problem-solving, is scarce. More research work should be conducted on the use of DGBL applications in relation to these skills, especially critical thinking, as it is one of the most relevant and sought-after skills in the competitive economy of the 21st century.

Furthermore, most of the DGBL applications in the teaching of science are not engineered based on specific guidelines and design theories. The review presented in this chapter covers numerous DGBL evaluation and design frameworks to determine the essential components needed to design and develop a DGBL framework that could help to enhance elementary students' critical thinking abilities.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

The primary goal of this research was to develop a DGBL framework and test its effectiveness using a game application. This chapter adopts the analysis, design, development, implementation, and evaluation (ADDIE) model, this model was used to limit mistakes throughout the design process and guarantee its robustness. The following sections provide a detailed discussion of every step of the ADDIE model.

3.2 Instructional Design

Since the proposed framework will be used for educational purposes, especially, to improve students critical thinking capacities, this research applied an instructional design approach to address the pedagogical aspects of the design process.

Molenda, Reigeluth, and Nelson (2003, p. 574) defined instructional design as ‘the principles and procedures by which instructional materials, lessons, and whole systems can be developed in a consistent and reliable fashion’. Meanwhile, Ritchey, Klein, and Tracey (2011, p. 3) characterised instructional design as ‘the science and art of creating detailed specifications for the development, evaluation, and maintenance of situations which facilitate learning and performance’.

In this study, the researcher applied a pedagogical design process to facilitate a fully focused approach that adopted an instructional design procedure. Such a technique paves the way for a simple step-by-step design process that helps to minimise mistakes and saves time during any phase of the development procedure (Eller, 2016).

3.3 The ADDIE Model

Among the various instructional design methods, the ADDIE model is considered to be one of the most prominent and popular (Shibley, Amaral, Shank, & Shibley, 2011).

The ADDIE model was constructed by the Centre for Educational Technology at Florida State University and first appeared in 1975 (Muruganantham, 2015).

This research employed the ADDIE model over other available models due to its adaptability and high flexibility, making it a convenient systematic problem-solving tool for instructional design in numerous fields, in addition, a considerable number of the newer instructional design models are based on the basic tenets of the ADDIE model (Davis, 2013; Hsu, Lee-Hsieh, Turton, & Cheng, 2014). Further, ADDIE provides instructional designers with a roadmap to the desired solution through practices rooted in iteration, revision, and step-by-step cyclical progression during the design process (Molenda, 2015; Branch, 2009; Bugis, 2018). Using the ADDIE model, the ICMDCR framework progressed from analysis to design, development, implementation, and evaluation. In the context of this research, the five phases of the ADDIE model were conducted as follows.

3.3.1 Analysis

The analysis phase is the foundation for all other phases of instructional design. This phase is primarily centred around the awareness of an emerging or existing problem (Muruganantham, 2015). In this research, recognition of the problem was met by considering the need to provide a DGBL framework to improve young learners' critical thinking skills.

A thorough literature review was conducted to acquire the prerequisite understanding of the problem. This review signified that students who learn via the traditional method of teaching demonstrate poor critical thinking skills. Meanwhile, research works that have utilised DGBL rarely focus on promoting students' 21st-century skills and particularly their critical thinking skills.

In addition to reviewing the literature, this study conducted needs analysis. According to Bugis, (2018) and Morrison, Ross, Morrison, and Kalman, (2019), needs analysis highlights the difference between what is and what needs to be, to determine whether an educational change is required and warranted (Cook & Ellaway, 2015). Information concerning needs analysis can be collected in a number of ways such as interviews, surveys, and formal assessments (Cook & Ellaway, 2015). Among these methods of information collection, interviews have been considered as a useful technique to obtain valuable information from the potential users of the proposed system (Lim, 2011; Pacheco, García, & Reyes, 2018). Therefore, in this research, needs analysis was conducted using interviews with teachers and primary-aged students.

The researcher held a few meetings with science teachers to get an insight into their thoughts concerning the traditional method of teaching, and, more importantly, the formulation of a DGBL framework and the creation of a gaming application. These meetings were brief and conducted in March 2018, (Interview questions, Appendix A).

When asked about the limitations presented by using the traditional method of teaching. Mrs. I, a science teacher in the school in which this study was conducted stated that this method of instruction is exhaustive as it is based entirely on the role of teachers in the teaching process. Mrs. N, also a science teacher in the school argued that in this method students' involvement is very limited due to a number of issues such as time constraints.

Concerning the potential impact of the traditional method, Mrs. I identified the traditional method of teaching as an approach that encourages memorisation of the learning content and inhibits students' exposure to other sources of information, while Mrs. N explained that in the traditional classroom, teachers are the main source of information and knowledge, while students are only recipients of this knowledge, hence, such practices does not allow students to develop their critical thinking.

Finally, concerning what the gaming application should focus on, Mrs. I explained that the gaming application needs to encourage students to improve their cognitive skills, especially, their critical thinking, whereas Mrs. N articulated that the gaming application is required to make science learning appealing to students.

With regards to learning characteristics and motivation, the school in which the study was conducted, and the majority of public schools in Malaysia rely mostly on the traditional teacher-centred learning approach. Hence, most students feel detached and isolated during the learning experience, as this teaching method neither engage students nor advance their critical thinking. To overcome these issues, this study attempt to provide students with an interactive DGBL application that improves their critical thinking while placing them at the centre of the learning experience.

3.3.2 Design

The second phase of the ADDIE model is based on the findings of the analysis phase (Muruganantham, 2015). During this phase, the researcher identifies the science units being taught, instructional objectives, learning strategies, and evaluation and testing methods (Muruganantham, 2015).

In this study, the proposed gaming application will cover three course units. First, energy transfer among living things, plants' characteristics, and energy characteristics. This study was conducted at the beginning of the school year and these units are the ones that the Malaysian science curriculum for fifth-graders covers during that time frame. Hence, the game was based on these three units.

The first instructional objective was for students to able to improve their critical thinking by using a gaming application called Ecoship Endeavour, which was based on the learning strategies and guidelines of the ICMDCR framework. The second instruction objective was for students to able to enhance their science subject knowledge by using the gaming application Ecoship Endeavour, which was based on the ICMDCR framework.

When learning via the proposed DGBL approach students will be interacting with different learning strategies such as information analysis and evaluation, decision-making, and discussion and collaboration with a teammate. These strategies were addressed in a more detailed manner in Section 2.15.

With regards to the evaluation methods, this study, utilised two instruments one for measuring students' critical thinking, while the other one was employed to measure student knowledge of the science subjects covered in the game (more information regarding these instruments is detailed in Section 4.7.2. In addition to these instruments, students while playing the game students need to solve a number of multiple-choice questions and puzzles related to the mission they just played.

A brief use case model was generated to illustrate the general layout of the proposed gaming application. This stage of research involved understanding and accommodating stakeholders' needs from the pedagogical dimension.

3.3.2.1 Use Case Model

Use cases are static descriptions that capture and communicate the expected behaviour of the software being developed (Klimek & Szwed, 2010; Munassar & Govardhan, 2011). Use cases are written from the perspective of the user as a flow of events. In this context, the user is recognised as an 'actor', and the narrative of the flow of events between this actor and the system is identified as the ‘use case’ (Klimek & Szwed, 2010; Oliveira et al., 2015). Use cases appear to be relatively easy to comprehend, even for people who are not familiar with IT, as use cases enable the understanding of the system without the need to delve too much into the implementation details (Cota et al., 2017). Figure 3.1 illustrates the use case scenario for Ecoship Endeavour. From this diagram, the science teachers were able to extract information effortlessly. For example, the teachers knew that Ecoship Endeavour would be played collaboratively in teams of two students and would include three gaming levels and the topics being covered. The teachers also understood the mechanics that would govern students’ progress from one level to the next.

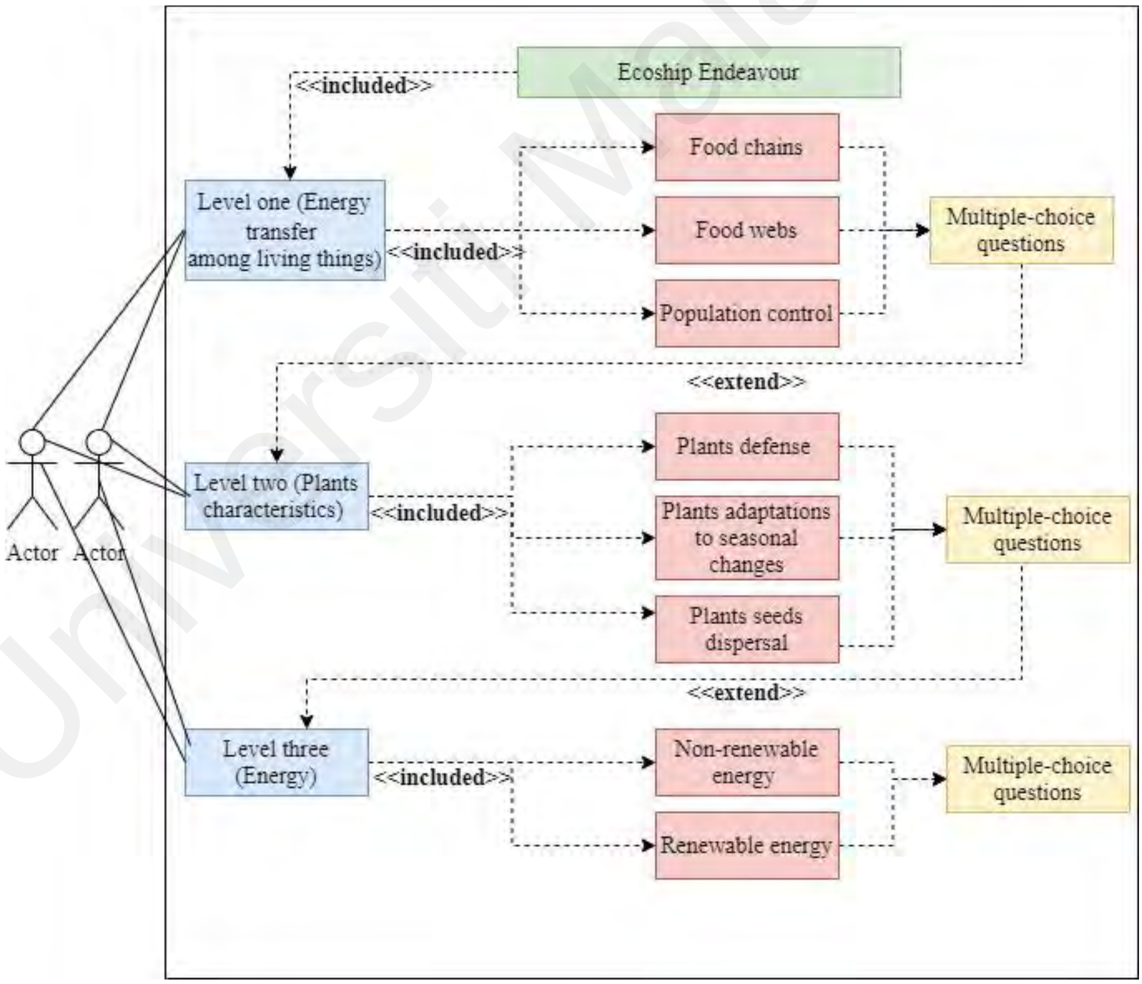

Figure 3.7: Ecoship Endeavour Use Case Diagram

3.3.3 Develop

The third phase of the ADDIE model is based on the suggestions and findings of the previous two phases: analysis and design. In this phase, the underdeveloped and provisional versions of the framework were transformed into a fully developed DGBL framework that could be embedded into the creation of a DGBL application, which could, in turn, be used by students in schools or educational institutes.

Therefore, this phase is mainly concerned with addressing the first research question: 'What are the most important components required when designing a DGBL framework to improve critical thinking?'.

A detailed description of the ICMDCR components was provided in Sections 2.15 and 2.16.

3.3.4 Implement

The fourth phase of the ADDIE model is the implementation phase. This phase is concerned with the effective and efficient delivery of learning instruction – whether classroom-based or computer-based – to improve students' understanding of the learning material (Muruganantham, 2015).

In this research, the implementation phase was based on the delivery of a DGBL application to students in Malaysian public schools.

The implementation phase of the ADDIE model involved an iterative development cycle similar to those used in software engineering; more detail about this process is given in Section 4.2. This phase addresses the second research question: 'How might one develop a DGBL application based on the proposed framework?'.

3.3.5 Evaluate

The final phase of the ADDIE model is the evaluation process, which focuses on testing the effectiveness of the learning material (Muruganantham, 2015). This phase was used to test the effectiveness of the proposed DGBL framework via Ecoship Endeavour in improving students' critical thinking skills.

This section addresses the third and final research question: 'How might one evaluate the effectiveness of the proposed framework using the DGBL game application?'.

In this research, the evaluation phase was performed twice: firstly, in the form of a pilot study (Section 4.6), and then the principal, quasi-experimental study (Section 4.7). Figure 3.2 below briefly illustrates how the ADDIE model was employed in this study.

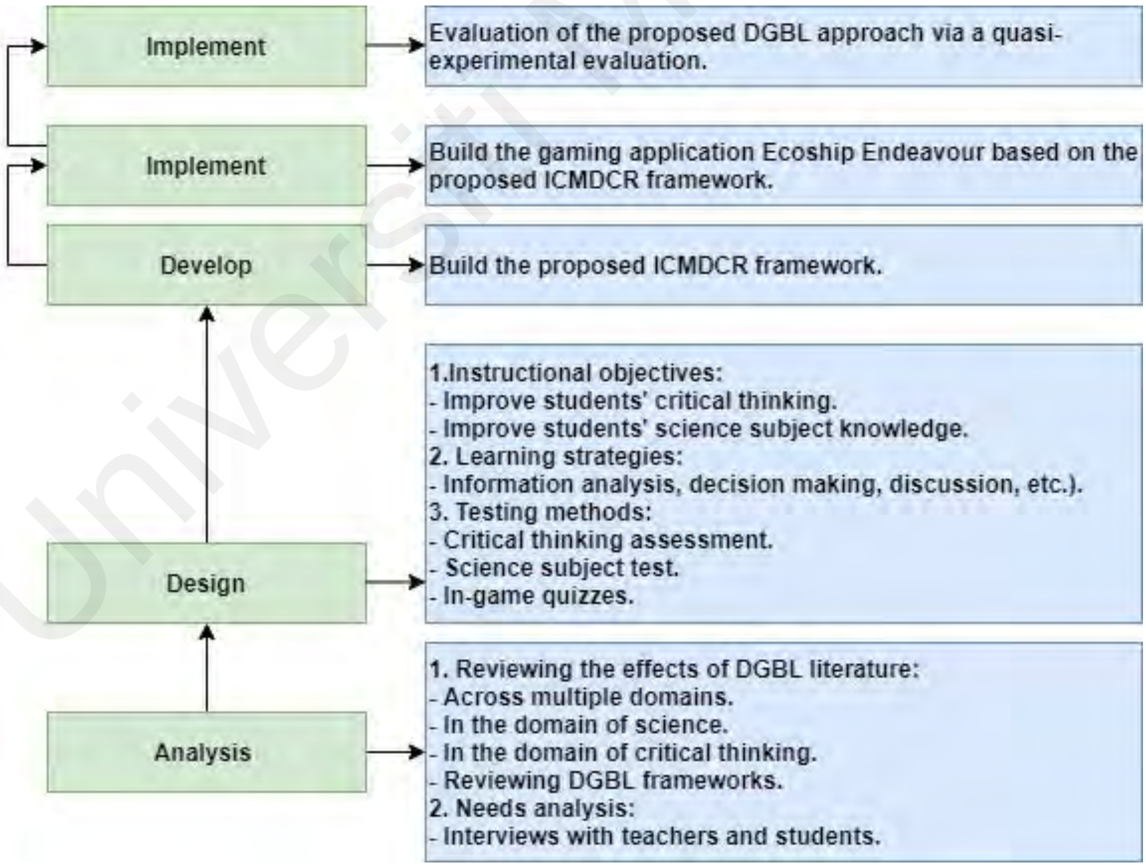

Figure 3.8: Summary of ADDIE Model Phases

3.4 Summary

In this study, an instructional design approach, namely, the ADDIE model was utilised to guide the research process from its early steps until the gaming application was evaluated by students in a formal classroom environment.

CHAPTER 4: DEVELOPMENT OF A DGBL APPLICATION: ECOSHIP ENDEAVOUR

4.1 Introduction

This chapter explains how Ecoship Endeavour was developed based on the proposed framework, its gaming content, how each level is played, and how components of the ICMDCR framework were integrated into the gaming environment. In addition, this chapter also explains the pilot study and the principal study (quasi-experimental) conducted for answering the third research question.

4.2 Development of a DGBL Application: Ecoship Endeavour

To test the hypotheses established in Chapter 1, the researcher designed and developed a computer game named Ecoship Endeavour to enhance students' critical thinking skills. Ecoship Endeavour was created using Construct 2, a powerful HTML5 game creator designed specifically for 2D game development. It allows anyone to build computer games without the need to have an extensive background in programming (Create Games with Construct 2, 2019). Since most public schools are not equipped with high-end IT infrastructure, Ecoship Endeavour was designed to consume limited processing and computing powers. To run Ecoship Endeavour smoothly, each laptop/PC should at least have 143.00 megabytes of storage, a 64-bit architecture, and a minimum of two gigabytes of random access memory.

The game was installed in Laptops provided by the researcher and science teacher.As shown in Table 4.1, a review of numerous very popular DGBL applications was carried out before Ecoship Endeavour was developed. Such a process offered a valuable insight into the manner in which such applications are developed. The reference games were selected from Commonsense.org. Common Sense is a non-profit organisation founded by James Steyer, an author, attorney, and Stanford University lecturer. The organisation aims to provide 'education and advocacy to families to promote safe technology and media for children' (Rutenberg, 2003).

Table 4.9: Popular Serious Games in Science

Game Name	Description	Cost
Marco Polo Ocean	Free play and puzzles combine to help students learn about marine life and vessels. This application is fun; it helps students gain an appreciation for the waters and the creatures that inhabit these worlds.	Free
Space	An interactive gaming application that allows students to learn about the sun, moon, and planets through interactive, open-ended play.	$2.99
Habitactics	This DGBL application allows students to discover and explore relationships in ecosystems.	$7.99
Crazy Plant Shop	This is a compelling visual style combined with clever integration of Punnett squares that makes learning genetics an attractive process that occurs naturally through play.	$5.99
Cell Strike	Thorough facts and information, a helpful tutorial, and intriguing gameplay will keep students trying to complete levels in order to learn more about the functions and components of the immune system.	Free to download; $7.99 unlocks the full game.
Who Wants to Live a Million Years?	In this game, Charles Darwin himself sets the scene for a hands-on introduction to natural selection. The game gives students a thoroughly interesting and interactive introduction to natural selection, paving the way for deeper study.	Free
Crazy Gears	This is a clear, simple, and fun introduction to science and engineering topics that may otherwise feel unapproachable for some students.	$2.99
Simple Machines	Students create and figure out simple machines and explore important physics concepts on their own terms.	$2.99
Inventioneers	Students gain experience in science and engineering practices through trial and error as they are seriously challenged to design, test, and redesign in a fun and playful setting.	$2.99
World of Goo	The game's design and the implicit knowledge embedded within this gaming application are encouraging factors to generate students' interest in geometry and the elements of building structures.	$19.99

Motion Force	Fun puzzles that make physics accessible and approachable. This gentle introduction to projectile motion will stick with students.	$5.99

Reviewing the games listed in Table 4.1 demonstrated that these games utilised captivating graphics. However, it also revealed that most of these games are not free and require a monetary subscription, hence, they could be inaccessible for students from economically disadvantaged families. However, this limitation is addressed in Ecoship Endeavour, as this gaming application requires no royalty fees from the schools or students using it, and this, in turn, could increase its accessibility.

In addition, some of these games mostly targeted general science topics. As such, they might not a viable option for schools, as schools seek DGBL applications that uniformly fit their curriculum. Further, there is a sense of vagueness regarding the learning principles that guided the development of these DGBL applications.

4.3 Prototyping Process

In this research, the Ecoship Endeavour prototype was developed based on the proposed ICMDCR framework to improve students' critical thinking skills and science knowledge. The prototyping process was conducted in accordance with the iterative design method. The iterative approach was used because it enables a researcher to perform short development cycles to ensure that the final product is tested and redeveloped as required based on stakeholders' feedback and comments (Crow, 2015). Commercial games are developed by hundreds of people, including user interface developers, game engine programmers, modellers, and animators of avatars and in-game creatures (De Lia & Fredericks, 2005; Rankin, 2008); they also receive substantial financial support and are usually granted a lengthy development process averaging 2–3 years (Farhan, 2015; Fullerton, Swain, & Hoffman, 2004). In contrast, most educational games are developed by researchers who frequently develop these games with very limited resources (Cheng et al., 2015; Hainey et al., 2016). Therefore, they run the risk of developing a game that might not meet the expectations of players who are familiar with high-quality commercial video games, which could negatively affect the gameplay experience and lead, in turn, to adverse learning outcomes (El-Nasr & Smith, 2006). This method permits the developer to identify the flaws in the system being developed and attend to them quickly; thus, new and improved software is produced after each iteration (Sarneabat & Kabir, 2018).

4.3.1 Placeholder Prototyping

The starting block for Ecoship Endeavour was a placeholder prototype. Placeholders can be defined as 'assets [that] are temporary resources used during the development of a game in place of final resources that haven't been created yet' (Vaillancourt & Egli, 2011, p. 275), which means that placeholders will ultimately be replaced by the final assets (Zagal & Altizer, 2015). Zagal and Altizer (2015) recommend that game developers use placeholders, arguing that it is an essential step in the game development process. Utilising this process allows the developer to reduce bottlenecks and helps different stakeholders to narrow their focus on specific aspects of the game. As such, they form a better comprehension of the nature of the game being developed.

In Ecoship Endeavour, placeholders included but were not limited to

- a placeholder for the loader splash screen (e.g., the loading screen),
- a placeholder for the dialogue system,
- placeholders for obstacles,
- placeholders for collectable items, and
- placeholders for level maps (such as forest, farms, and cities).

4.3.2 Nature of the Iterative Cycles

The iterative cycle involved interacting with the same three science teachers who shared their input on Ecoship Endeavour and communicated students' suggestions. In total, there were four iterative cycles for each level. Tables 4.2 – 4.5 demonstrate the early cycles of the iteration process of Ecoship Endeavour's first level. These cycles were concerned with seeking feedback from the science teachers, while the later stages of the prototyping process sought input from both teachers and students.

Table 4.10: Iteration One

Mode of Testing
Placeholder prototype tested with other teachers
Discussion concerning the potential directions for development
Two-dimensional game world for a single level
Basic user interface (loading screen, saving mechanism)
Avatars (the in-game playable characters)
Game mechanics developed
2D platformer obstacles, such as platforms and walls
Feedback from teachers
All dialogue in the game should be in one language

Table 4.11: Iteration Two

Mode of Testing
Improved prototype with more details evaluated by science teachers
Discussion about potential directions for development and the nature of science learning behind the gaming tasks
Features tested in iteration two
The mechanism of food chains
The mechanism of food webs
The impact of one animal on the other animals
Feedback from teachers
All three topics in this subject should reuse the same game systems, simplifying the design and thereby allowing the student to focus upon the lessons without having to keep learning new rule sets, which could undermine the learning process

Table 4.12: Iteration Three

Mode of Testing
The mechanism of food chains
The mechanism of food webs
The impact of one animal on the rest of the animals
Features tested in iteration three

Identify the producers and consumers of certain locations (e.g., ponds, forests, and paddy fields)
Identify the hierarchy of animals in certain locations (e.g., ponds, forests, and paddy fields)
Feedback from teachers
Simplify the locations and make them smaller to prevent students from wandering in any given area, which could distract the student, induce boredom, and demotivate them

Table 4.13: Iteration Four

Mode of Testing
Four students tested the fully functioning prototype of level 1
Science teachers tested the fully functioning prototype of level 1
Features tested in iteration four
Experimenting with food chains
Experimenting with food webs
Experimenting with the impact of one animal on the rest of the animal population
Feedback from students and teachers
Teachers reported that students had a noticeable sense of enjoyment when they played the game
Teachers also reported that students expressed their desire to include more levels and expand the game
Teachers suggested that a Malay language expert should be hired to check the language and ensure that all the text used in the game is grammatically correct and free of error

4.4 Contents of Ecoship Endeavour

The game covers various science concepts, allowing students to gain knowledge and solve puzzles and multiple-choice questions before advancing to the next level of the game that addresses another science topic.

The storyline of the game features two alien robots who visit Earth to teach young children that, if they wish to protect the Earth for future generations, they are required to improve their science knowledge. The science topics taught by the alien robots are: (a) energy transfer among living things, such as food chains, food webs, and how the population of animals in a food web is based on other animals; (b) plant characteristics, such as protection, survival, and seed dispersal; and (c) energy, both renewable and non-renewable. These three topics are covered in the standard five Malaysian science subject syllabus.

4.4.1 The Welcome Screen

The welcome screen aimed to allow students to learn about and familiarise themselves with the gaming application without external guidance. Such a step is essential to keep students focused on the learning process and not overwhelm them with an excessive overflow of information. Therefore, the interface of Ecoship Endeavour was intended to look uncomplicated, straightforward, and prevent students from spending the majority of their times time learning how to operate the game, as illustrated in Figure 4.1.

Figure 4.9: The Welcome Screen

4.4.2 Gaming Content

Ecoship Endeavour is an educational game developed specifically for this study. In this game, students were provided with missions to explore; however, before engaging in the playing process, students were briefed on the nature of the mission that awaited them. The game is organised in a series of tasks where students were required to collect items, recover in-game clues, solve puzzles, overcome challenges, and collaborate with teammates.

The gaming content of Ecoship Endeavour along with its instructions and messages were delivered to students in Bahasa Melayu. Instructors stated that presenting the gaming application to students in their native language would be more appropriate, as it would ensure better interactions with the game, which may result in an improved learning experience. However, two identical applications were developed one in the Bahasa language and the other one in English. The English application was used solely by the researcher, as the researcher was not a native speaker of the Malay language.

In the first level of Ecoship Endeavour, the students explored a factory location, as shown in Figure 4.2 (A); they also investigated a forest location, as shown in Figure 4.2 (B). The first gaming mission aimed to teach students about energy transfer.

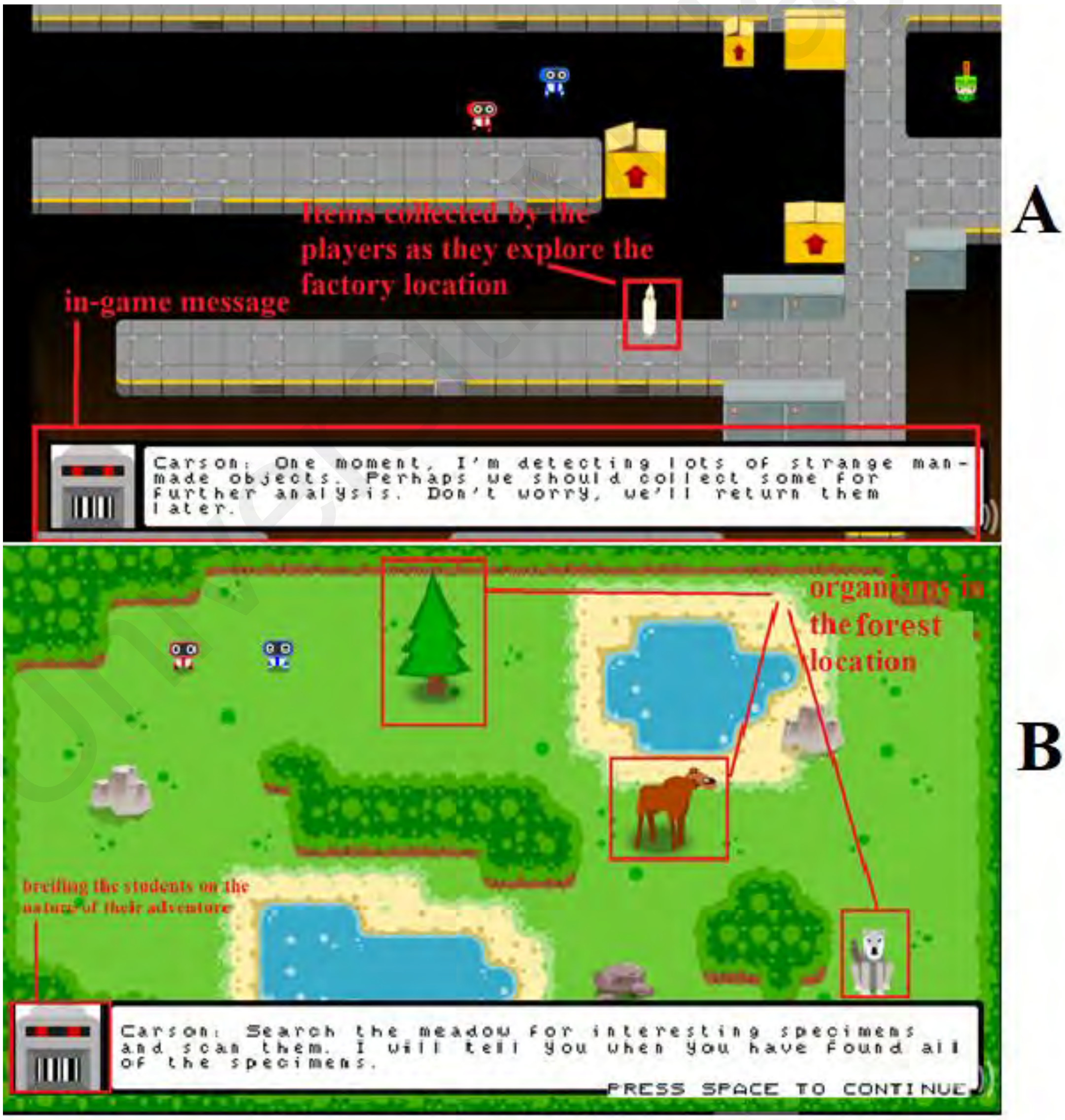

Figure 4.10: Screenshots from Ecoship Endeavour (Level One)

When students advanced to the next level, they learned about plants' characteristics. More specifically, this level attempted to bring the students' attention to plants' protection, adjustments to seasonal changes, seed dispersal methods, and the importance of these features in plants' survival. In this level, students examined two contrasting environments. In the first, they investigated, analysed, and interviewed farmers assuming the role of non-playable characters whose plants have lost their ability to disperse seeds and, as such, can no longer survive, as illustrated in Figure 4.3 (A). Conversely, Figure 4.3 (B) shows students inspecting a healthy farm in which the plants are capable of dispersing their seeds and can thus ensure their survival.

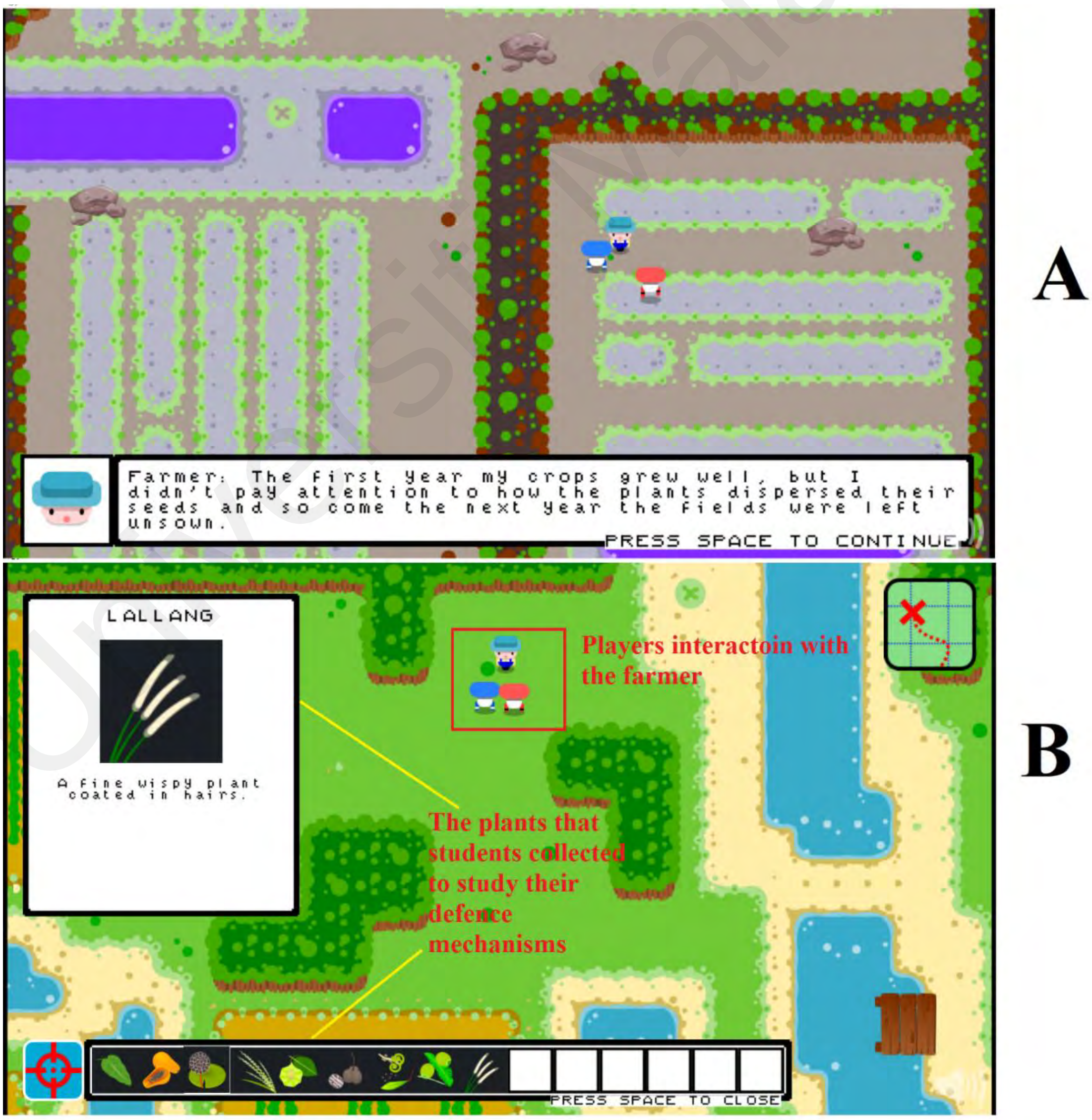

Figure 4.11: Screenshots from Ecoship Endeavour (Level Two)

Finally, in the third level, students learned and recognised the difference between renewable and non-renewable energy by investigating and assessing the effects of different energy types on people's lives, as shown in Figure 4.4.

Figure 12.4: Players in the Hardware Store Collecting Various Energy Resources

4.5 Embedding ICMDCR in Ecoship Endeavour

4.5.1 Inquiry

In Ecoship Endeavour, players can interact with inquiry-based learning activities multiple times throughout their learning experience. For example, in level one, the students investigated the cause-and-effect relationships among organisms in a particular food chain or food web. This investigation was implemented by adjusting an increase or decrease dial, which could help students interpret and recognise the impact of any surge or decline of one organism on the rest of the population in that food web.

In the second level, students explored several locations (e.g., farms and gardens) and collected plants. Students collected these plants to learn about their characteristics, such as their defence mechanisms. To recognise the distinctive defence properties of these plants, the students needed to solve a puzzle whereby they were required to analyse the information associated with each plant and defence mechanism to infer which plant adopts which defence technique. A similar learning activity was applied to other plant characteristics (e.g., seed dispersal and weather adaption techniques).

In the third level, students navigated a city suffering from energy-related problems (e.g., power outages and insufficient fuel for heating). To comprehend the nature of these issues, students interviewed the city's inhabitants and experimented with two types of energy (renewable and non-renewable) to recognise the difference between the different forms of energy and discern the unique attributes of both types.

4.5.2 Communication

As mentioned in Section 2.16.2, communication in Ecoship Endeavour was manifested in two ways. Firstly, communication occurred between team members, allowing them to exchange information and discuss the direction of the game, which helped students to approach the challenge presented to them from multiple angles, resulting in better problem-solving. The second type of communication took the form of a cue, a tip, or feedback, which students received when they made a mistake or sought clarification during gameplay when facing a challenge.

4.5.3 Mystery

To sustain players' engagement and curiosity, Ecoship Endeavour introduces elements of mystery throughout the game. The students could experience these elements as they played and interacted with the non-linear storyline of the game, providing them with different perspectives and scenarios of the topic they were exploring.

4.5.4 Decision Making

Students interacted with the decision-making feature at the beginning of each level, where the game highlights some of the details pertaining to the mission and its goals. For example, at the start of level two, students were provided with a summary in relation to plants and their characteristics. A section of the briefing session aimed to bring students' attention to why plants have developed defence mechanisms. At this point, the students were given a choice to explore a farm where plants cannot defend their enemies or a farm where plants have retained their ability to protect themselves against predators. However, before advancing to the mission, students were required to analyse the information they received during the briefing session and review its impact on the gaming process.

It is worth noting that Ecoship Endeavour was designed for educational purposes. As such, when students made the wrong decision, they received no penalty or reduction of points; instead, they were asked to assess and examine why a particular decision was incorrect before being redirected to the right decision.

4.5.5 Challenge

In this gaming application, challenges were incorporated to motivate the students to overcome obstacles encountered as they played. The presence of such barriers served to improve students' immersion in the game and promote their problem-solving capabilities. In Ecoship Endeavour, challenges fall into two main categories. Firstly, incremental challenges can be observed in that the difficulty of the game increases as players advance from one level to the next. Secondly, the game incorporates logical challenges that examine the players' abilities to answer multiple-choice questions and solve puzzles.

4.5.6 Rewards

Ecoship Endeavour implemented a reward system that recognised students' accomplishments and the progress achieved through their gaming journey. In this DGBL application, rewards were manifested when a student answered a question or solved a puzzle. The student would receive three points, resulting in an incremental change in the progress bar measuring the student's advancements and a positive encouragement message.

4.6 Pilot Study

The pilot study was designed to identify any weakness in the research design and was performed under the same conditions and with the same instruments as those intended for the principal study. The main reason behind conducting the pilot study was to ascertain if the design could be applied effectively, as well as to detect errors in the data collection and interpretation processes.

4.6.1 Content Validity

In order to improve the validity of this assessment, the critical thinking test was reviewed by a panel of five experts – three science teachers and two academics – who were familiar with science teaching, e-learning, and educational psychology. In this research, the reviewing process was utilised to discover potential flaws and issues in the assessment, ensure the robustness of the scores' interpretations, and assess the extent to which these scores measured what they were designed to measure.

4.6.2 Instrument Reliability

The piloting of the instrument was conducted among 30 respondents to test the reliability of the critical thinking test. This study implemented the Kuder–Richardson-20 test. The Kuder–Richardson-20 is a statistical measure that aims to compute the internal consistency of dichotomous scales (DiIorio, 2006). A Kuder–Richardson coefficient value of 0.70 or higher indicates a satisfactory level of internal consistency. Table 4.6 shows that all coefficient values were between 0.723 and 0.777, thereby surpassing the recommended value of 0.70.

Table 4.14: Kuder–Richardson Coefficient Values for the Research Instrument

Scale	Coefficient value
Hypothesis identification	0.740
Induction	0.723
Deduction	0.735
Explanation	0.777
Evaluation	0.724

4.6.3 Test-Retest Reliability

Intraclass correlation coefficient (ICC) is a widely utilised reliability index in test-retest or repeatability procedures, which demonstrate the degree of correlation between at least two quantitative measurements (Koo, & Li, 2016; Perinetti, 2018). According to Koo and Li (2016), ICC values lower than 0.50 indicate poor reliability, values that range between 0.5 and 0.75 suggest moderate reliability, values between 0.75 and 0.9 suggest good reliability, and values greater than 0.90 indicate excellent reliability. ICC estimates were calculated at the 95% confidence level based on a 2-way mixed-effects model, all ICC values for the critical thinking test were between 0.703 and 1.00, which indicates moderate to excellent intrarater reliability between pre- and post-tests for the critical thinking test. As seen in Table 4.7.

Table 4.15: ICC Coefficients for the Critical Thinking Test

Item	ICC coefficients	95% Confidence Interval		Sig
		Lower Bound	Upper Bound	
Hypothesis identification 1	0.895	0.736	0.959	0
Hypothesis identification 2	0.889	0.719	0.956	0
Hypothesis identification 3	0.888	0.716	0.956	0
Hypothesis identification 4	1.000	1	1	.
Induction 1	1.000	1	1	.
Induction 2	0.936	0.839	0.975	0
Induction 3	1.000	1	1	.
Induction 4	0.878	0.693	0.952	0
Deduction1	0.927	0.815	0.971	0
Deduction 2	0.936	0.839	0.975	0
Deduction 3	0.862	0.65	0.945	0
Deduction 4	0.846	0.611	0.939	0
Explanation 1	0.703	0.25	0.882	0.006

Explanation 2	0.813	0.527	0.926	0
Explanation 3	0.772	0.424	0.91	0.001
Explanation 4	1.000	1	1	.
Evaluation 1	0.950	0.873	0.98	0
Evaluation 2	0.878	0.693	0.952	0
Evaluation 3	0.936	0.839	0.975	0
Evaluation 4	1.000	1	1	.

4.7 The Principal Investigation

This section details how the experimental procedure of this study was conducted.

4.7.1 Participants

The students who participated in this study were fifth-grade students aged 11 from a primary school located in Kajang. As the participants were below 18 years old, a consent letter was obtained from their respective parents to allow them to participate in the study (Consent form, Appendix B). The students' participation was voluntary, and their identities were kept confidential; furthermore, the students were informed that they could withdraw at any given moment without being penalised on their science subject marks.

In total, 124 students from four Grade 5 classes participated in this study. These students were divided equally between two groups: 62 students were assigned to the experimental group, while the other 62 students were assigned to the control group. The total number of participants is sufficient for conducting this study, as Sekaran and Bougie (2016) stated that a research sample between 30 and 500 is satisfactory for scientific research.

Female students from different ethnic backgrounds (e.g., Malay, Indian, and Chinese) and medium socioeconomic backgrounds participated in this study. It is worth mentioning that all the students who participated in this study were female students, as the school in which the study was carried out was an all-girls school.

With regards to the sampling mechanism, this study employed the convenience sampling approach. This sampling approach was utilised to recruit all the students who participated in the study, employing this approach is justified on the basis that the researcher needs to select public schools based on their availability and interest to partake in this study.

4.7.2 Instrument

The pre- and post-tests of critical thinking were based on the critical thinking skills test introduced by Yeh (2003, 2009, 2012), which has a Cronbach's alpha value of 0.80 and is based on the Cornell critical thinking test (Ennis, Millman, & Tomko, 1985) and the Watson–Glaser critical thinking test (Watson & Glaser, 1980).

The test comprised 20 multiple-choice questions, distributed evenly according to five dimensions. These dimensions were hypothesis identification, which refers to students' ability to highlight the underlying ideas or indirect assumptions of the problem being investigated; induction, which is primarily concerned with the capacity to establish a connection between a set of specific examples to build a reasonable generalisation of the issue at hand; deduction, which is centred around students' capability to draw logical conclusions from a set of general statements to build a rational conclusion of the issue being addressed; explanation, which refers to students' ability to determine which phenomena or casual relationships are implied by the given statement; and evaluation of arguments, which reflects students' ability to assess the strength of an argument (Kong, 2015; Yang & Chang, 2013). Students answered the questions individually and were given 20 minutes to solve the questions (1 minute per question). Each question consisted of a statement followed by three multiple-choice answers. The total possible score of the test was 20 points; thus, a valid response was awarded one point, and an invalid answer was assigned zero points (for a sample set of questions, see Appendix C). Hence, a higher score reflects better critical thinking capabilities. The validation and reliability of this instrument are addressed in Section 4.6.

With regard to the science subject test, the test was designed and marked by the science teacher, included 20 items, and comprised multiple-choice items and questions that focused on students' abilities to memorise information, A correct answer was awarded five points, and, in a similar manner to the previous test, a higher score suggests better performance (for a sample set of questions, see Appendix D). Students answered all the questions individually and were given 40 minutes to complete the test. The total possible score of the test was 100 points.

4.7.3 Procedure

As shown in Figure 4.5, a primary school was selected to conduct an experimental study for this research. After obtaining the appropriate forms from the headmistress of the school, an application to conduct a research study in this school was submitted to the MOE. Once approval had been received from the MOE, the researcher and science teachers identified which science modules should be included in Ecoship Endeavour.

RCTs are perceived as the most rigorous form of research methods and often described as the 'gold standard' for conducting research (Ginsburg & Smith, 2016). Nonetheless, it is often challenging to follow the RCT design in educational settings when DGBL is involved because researchers are not always able to find an appropriate condition to be used as the control groups (Perrotta, Featherstone, Aston, & Houghton, 2013).

In addition, the quasi-experimental design was chosen because it was challenging to receive cooperation from the school administrators to perform complete randomisation of students due to administrative constraints, despite the fact that this study was permitted by the MOE, which is the governing body that regulates and governs academic research in Malaysian schools.

Before the learning activity, the teacher introduced the researcher to the students, and the researcher explained the necessary steps to the students, which entailed undertaking pre-tests and post-tests to evaluate their critical thinking and science learning performance.

During the learning activity, students in classes A and B were allocated to the experimental group, where they used Ecoship Endeavour, it is worth mentioning that students in the experimental group were taught by the DGBL application only and did not attend the regular science classes. Students in classes C and D were assigned to the control group, where they were taught using the traditional teaching method, where they listened to the science teacher and occasionally participated in paper-based drill and practice exercises. Hence, all the learning activities afforded by the ICMDCR framework such as discussions and interactions with teammates, instant feedback in case of a problem or a mistake during the learning process, and information evaluation and decision making were not available for students who were allocated to the control group.

It also should be noted that students in the control group could not access or use Ecoship Endeavour during this study, as the game was played on laptops provided by the researcher and the science teacher.

In the first week of the intervention, all students sat for the critical thinking and science achievement pre-tests; both tests were paper-based and lasted for 1 hour. After the tests, the researcher met the students of the experimental group to explain and highlight the features of the game and its functions.

Over the next three weeks, students in the experimental group played Ecoship Endeavour on laptops in teams of two, twice a week for six sessions. Each session lasted for 40 minutes. Meanwhile, students in the control group received their instruction via the traditional method of teaching. It should be noted that although students were guided separately under two different learning methods, the content of both methods was the same, and the four classes were taught by the same science teacher, who has more than 15 years' experience teaching science for primary school students.

By the end of the learning activity in the fourth week, students in the experimental and control groups again took the post-tests. These tests were also paper-based and aimed to assess students' critical thinking and science learning performance.

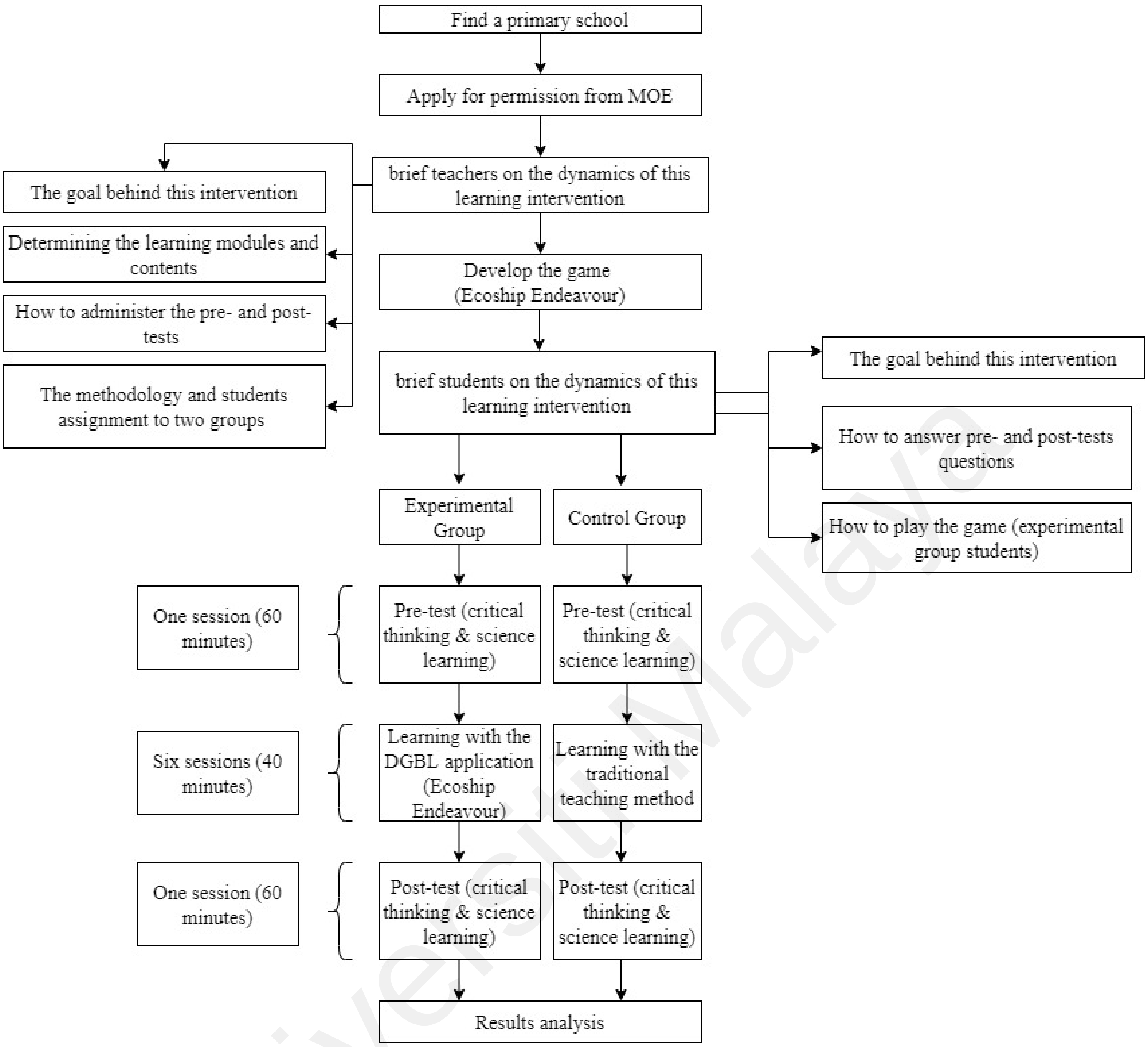

Figure 4.13: Research Procedure and Experimental Design

4.8 Summary

The chapter explained the mapping of the proposed ICMDCR framework and how this framework was embedded within the DGBL application to enhance students' critical thinking skills. In addition, this chapter also highlighted the procedures to administer and conduct a quasi-experimental study for this research. Further, Ecoship Endeavour was subsequently utilised in the evaluation of the ICMDCR framework through a quasi-experimental intervention on a sample of Malaysian students, which is detailed in the next chapter.

CHAPTER 5: ANALYSIS OF RESULTS OF QUASI-EXPERIMENTAL STUDY

5.1 Introduction

This chapter presents the analysis of data collected during the quasi-experimental study to answer the third research question. The data were analysed using Statistical Package for the Social Sciences (version 25). The results presented in this chapter include those of statistical tests (normality and independent *t*-tests). Furthermore, the researcher performed analysis of variance (ANOVA) and Bonferroni tests to ascertain if there were changes in students' results before and after the intervention in terms of their critical thinking abilities and science subject scores.

5.2 Normality Test

Distribution of data is a basic assumption in measuring the variation of variables is mentioned as normality in statistics. Normality of data can be measured by statistical methods, such as the kurtosis and skewness test, and graphical techniques, such as boxplots and histograms.

5.2.1 Skewness and Kurtosis

Skewness is a statistical procedure employed to characterise the degree of asymmetry of a given distribution, the skewness value of a perfectly distributed sample is zero. A positive skewness value shows that the tail is on the right side of the distribution. Conversely, a negative skewness value indicated that the tail is on the left side of the distribution (Kim, 2013). On the other hand, kurtosis is a measure of the peakedness of a distribution, the kurtosis value of a perfectly distributed sample is zero, a positive kurtosis value indicates high peaks, while a negative kurtosis value indicates a flat-topped curve (Kim, 2013).

For the control group of this study, the skewness test results were within the range of 0.57 and -0.357; for the experimental group, the skewness results were between 0.206 and -0.556. The kurtosis test values for the control group were between -0.067 and -1.128, while the kurtosis values for the experimental group ranged between 0.598 and -0.947. Based on the findings (shown in Table 5.1), all the pre-test variables were within the range of ±2. Hence, it can be concluded that all the variables were distributed normally (George & Mallery, 2019; Liu, Wang, & Koehler, 2019). In addition, the ratios of skewness and kurtosis to standard error (Std.Error) for both the control and experimental groups were within the range of ±3.29. It can, therefore, be inferred that all variables were distributed normally (Kim, 2013).

Table 5.16: Skewness and Kurtosis Values of All Research Variables (Pre-Test)

Variable	**Group**	**Skewness**	**Std.Error**	**Skewness/ Std.Error**	**Kurtosis**	**Std.Error**	**Kurtosis /Std.Error**
Hypothesis identification	Control	0.57	0.304	1.88	-0.304	0.599	-0.51
	Experimental	0.206	0.304	0.68	-0.259	0.599	-0.43
Induction	Control	-0.043	0.304	-0.14	-1.128	0.599	-1.88
	Experimental	0.011	0.304	0.04	-0.947	0.599	-1.58
Deduction	Control	-0.326	0.304	-1.07	-0.067	0.599	-0.11
	Experimental	-0.182	0.304	-0.60	-0.562	0.599	-0.94
Explanation	Control	-0.357	0.304	-1.17	-0.713	0.599	-1.19
	Experimental	-0.556	0.304	-1.83	0.598	0.599	1.00
Evaluation	Control	0.194	0.304	0.64	-0.456	0.599	-0.76
	Experimental	-0.182	0.304	-0.60	-0.206	0.599	-0.34
Overall critical thinking	Control	0.162	0.304	0.53	-0.161	0.599	-0.27

	Experimental	-0.24	0.304	-0.79	-0.886	0.599	-1.48
Science subject test	Control	-0.057	0.304	-0.19	-0.49	0.599	-0.82
	Experimental	0.042	0.304	0.14	-0.304	0.599	-0.51

With regard to post-test results, the skewness test values for students in the control group were between 0.06 and -0.458, whereas the values for the experimental group were between 0.038 and -0.59. The control group's kurtosis results ranged between -0.164 and -0.86, while the experimental group values were between -0.465 and -1.249. According to these findings (see Table 5.2), all the pre-test variables were within the range of ±2, As Table 5.2 shows. Thus, according to George and Mallery (2019) and Liu et al. (2019), it can be concluded that all the variables were distributed normally. Moreover, the ratios of skewness and kurtosis to Std.Error for both groups were within the range of ±3.29. As such, these findings indicate that all variables were distributed normally (Kim, 2013).

Table 5.17: Skewness and Kurtosis Values of All Research Variables (Post-Test)

Variable	Group	Skewness	Std.Error	Skewness/ Std.Error	Kurtosis	Std.Error	Kurtosis /Std.Error
Hypothesis identification	Control	-0.25	0.304	-0.82	-0.454	0.599	-0.76
	Experimental	0.038	0.304	0.13	-0.465	0.599	-0.78
Induction	Control	-0.049	0.304	-0.16	-0.48	0.599	-0.80
	Experimental	-0.59	0.304	-1.94	-1.249	0.599	-2.09
Deduction	Control	-0.117	0.304	-0.38	-0.86	0.599	-1.44
	Experimental	-0.582	0.304	-1.91	-0.775	0.599	-1.29
Explanation	Control	-0.458	0.304	-1.51	-0.398	0.599	-0.66
	Experimental	-0.504	0.304	-1.66	1.113	0.599	1.86
Evaluation	Control	0.031	0.304	0.10	-0.217	0.599	-0.36

	Experimental	0.078	0.304	0.26	-0.757	0.599	-1.26
Overall critical thinking	Control	-0.365	0.304	-1.20	-0.342	0.599	-0.57
	Experimental	-0.184	0.304	-0.61	-0.661	0.599	-1.10
Science subject test	Control	0.06	0.304	0.20	-0.164	-0.304	-0.27
	Experimental	0.144	0.304	0.47	-0.813	-0.304	-1.36

5.2.2 Boxplots

Boxplots are an effective exploratory data analysis technique (Cooksey, 2020). Boxplots are utilised to provide a visual examination of outliers, as their presence can severely distort the data and result in a lack of normal distribution (Leys, Klein, Dominicy, & Ley, 2018; Mowbray, Fox-Wasylyshyn, & El-Masri, 2019). In this study, both groups in the pre- and post-test had no outliers regarding critical thinking, as shown in Figure 5.1. Similarly, there were no outliers in students' responses to the science subject in the pre- and post-tests (see Figure 5.2).

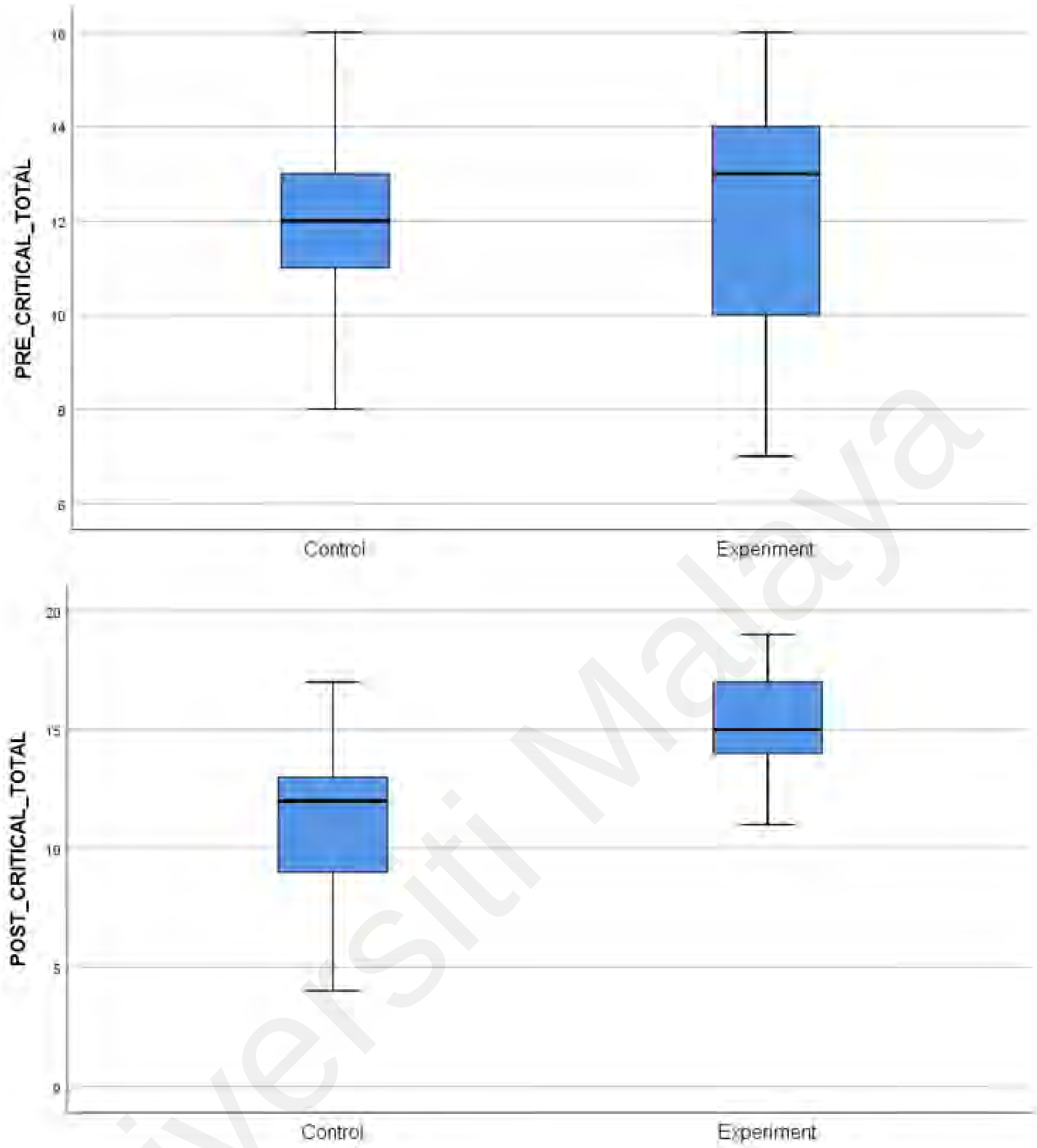

Figure 5.14: Boxplots for Critical Thinking in Both Groups Across Time

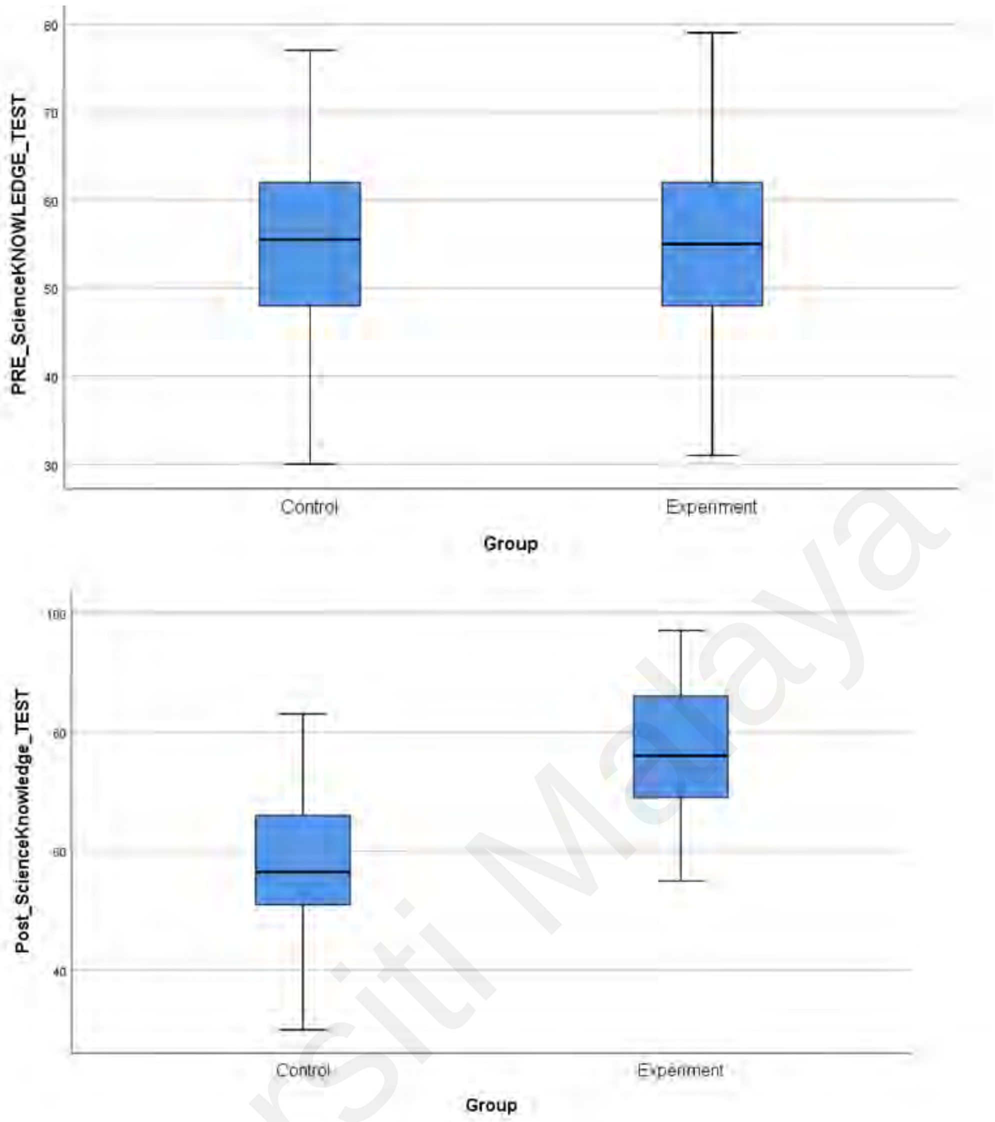

Figure 5.15: Boxplots for Science Subject in Both Groups Across Time

5.2.3 Histograms

Histograms are helpful tools as they provide a clear visual representation of data distribution. As Figure 5.3 shows, students' scores in the critical thinking test from both groups had a central distribution and fit appropriately within the bell, thereby implying normal distribution of data at the pre- and post-test levels.

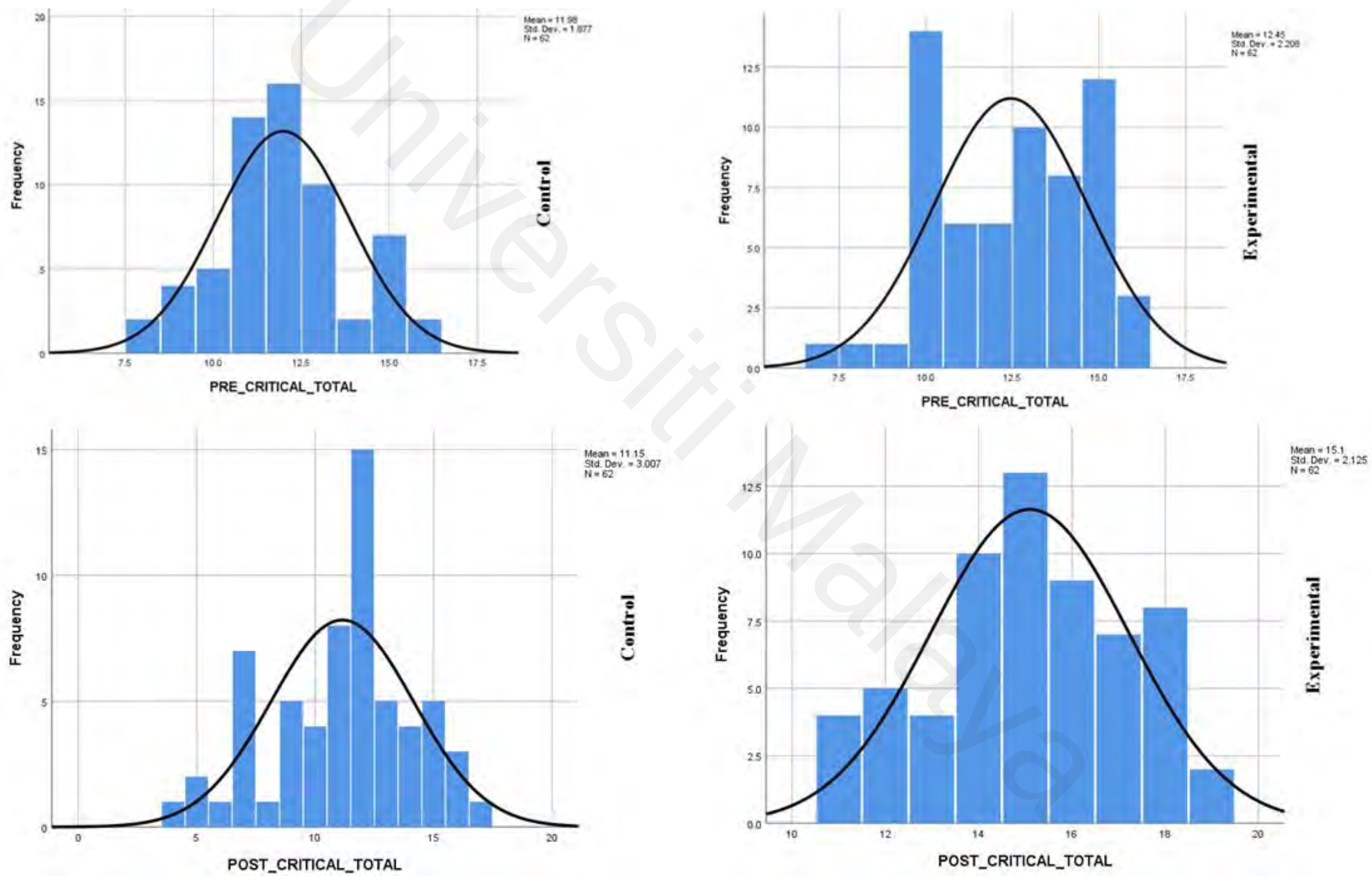

Figure 5.16: Histograms of the Critical Thinking Test

Concerning students' performance in the science subject test, Figure 5.4 shows that both groups had a central distribution of data. As such, it can be concluded that the data is normally distributed for both the pre- and post-tests.

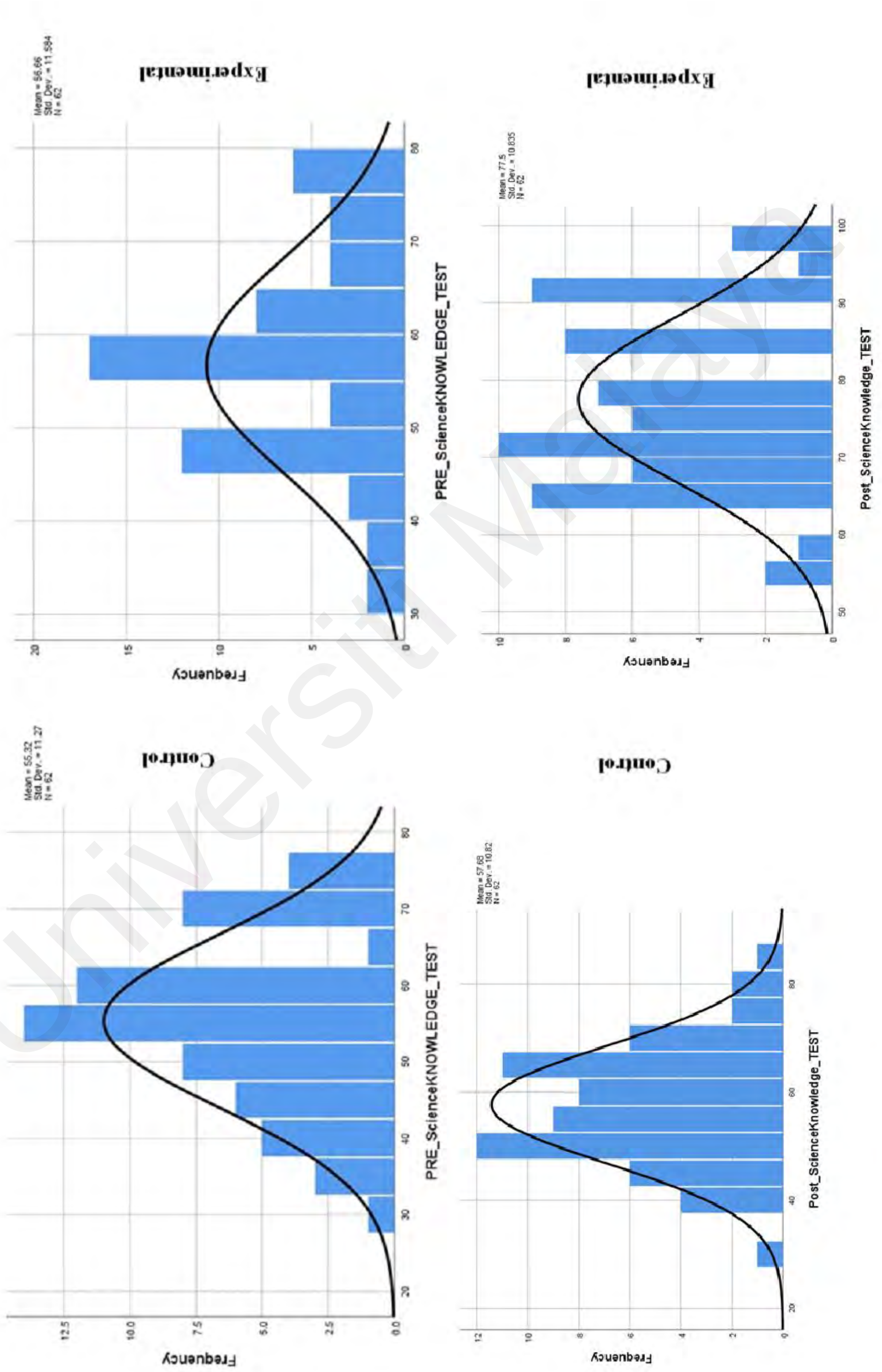

Figure 5.17: Histograms of the Science Subject Test

5.3 Homogeneity Analysis

Scholars recommend that prior to statistical data analysis, one should examine the assumption of homogeneity of the experimental and control groups. Independent sample *t*-tests were thus implemented separately for the critical thinking test and its subscales and the science subject test to address this issue. As Table 5.3 shows, the *p* value for the critical thinking test, every subscale of the critical thinking test, and the science subject test in both groups for the pre-test was above the significance threshold of 0.05, indicating no significant difference between students' performances, regardless of their group. The much higher values of the mean and standard deviation scores for "overall critical thinking" and "science subject test" compared to those of the individual critical thinking subscale (such as induction, deduction, etc.) is because the total score for each critical thinking subscale is only four marks, whereas the total score for the "overall critical thinking" test is 20 marks, and the total score for the "science subject test" is 100 marks.

Table 5.18: Mean Comparison Between Groups for all Research Variables at Pre-Test Level

Variable	Group	N	Mean	Std. Deviation	T value	P value
Hypothesis Identification	Control	62	2.02	0.878	0.86	0.392
	Experimental	62	1.89	0.791		
Induction	Control	62	2.53	1.036	-1.404	0.163
	Experimental	62	2.77	0.876		
Deduction	Control	62	2.23	0.895	-1.941	0.055
	Experimental	62	2.53	0.863		
Explanation	Control	62	2.95	0.876	1.339	0.183
	Experimental	62	2.74	0.867		
Evaluation	Control	62	2.26	0.828	-1.828	0.07
	Experimental	62	2.52	0.741		
Overall critical thinking	Control	62	11.98	1.877	-1.271	0.206
	Experimental	62	12.45	2.208		
Science subject test	Control	62	55.32	11.27	-0.652	0.515
	Experimental	62	56.66	11.584		

ANOVA Analysis

The following sections describe the application of a two-way repeated-measures ANOVA analysis to evaluate whether there were significant differences between the control and experimental groups in the critical thinking test and its subscales and the science subject test.

5.4.1 Critical Thinking Test

The following sections examine the hypotheses established in Chapter 1, Section 1.6. Specifically, these sections aim to review the following research hypotheses:

H_0: The gaming application which is developed based on the proposed DGBL framework does not enhance students' critical thinking skills.

H_{1a}: The gaming application which is developed based on the proposed DGBL framework does enhance students' critical thinking skills.

5.4.1.1 Hypothesis Identification

The findings illustrated in Table 5.4 show the descriptive statistics (*M* and *SD*) of the hypothesis identification subscale in pre and post-test. The pre- and post-test scores of the experimental group students in the hypothesis identification subscale were 1.89 (*SD* = 0.791) in the pre-test and 2.52 (*SD* = 0.825) in the post-test, while the mean scores of the control group were 2.02 (*SD* = 0.878) in the pre-test and 1.98 (*SD* = 0.896) in the post-test.

Table 5.19: Descriptive Statistics of Hypothesis Identification in Both Groups

Test	Group	N	Mean	SD
Pre test	Control	62	2.02	0.878
	Experimental	62	1.89	0.791
Post test	Control	62	1.98	0.896

	Experimental	62	2.52	0.825

As shown in Table 5.5, the findings for the within-subjects effect (time) of the repeated-measures ANOVA were significant ($F_{(1,122)} = 7.635$, $p = 0.007$, $\eta^2 = 0.059$). The main effect of group was not statistically significant ($F_{(1,122)} = 3.121$, $p = 0.063$, $\eta^2 = 0.028$). These results indicate that the interaction between group and time was statistically significant ($F_{(1,122)} = 9.367$, $p = 0.003$, $\eta^2 = 0.071$), which suggests that the change of hypothesis identification was significantly different across the times (i.e., pre- and post-test).

Table 5.20: Summary of Repeated-Measures ANOVA for Hypothesis Identification Subscale

Source of variation	**Df**	**MS**	**F**	**P Value**	**η2**
Time	1	5.52	7.635	0.007	0.059
Group	1	2.52	3.515	0.063	0.028
Time * Group	1	6.778	9.376	0.003	0.071

To test the difference, the researcher applied a post-hoc test (Bonferroni) to compare the mean scores of hypothesis identification. According to the result of the Bonferroni test (see Table 5.6), the difference of the hypothesis identification mean scores between pre-test and post-test in the control group was not statistically different ($p = 0.833$), whereas this difference was statistically significant in the experimental group ($p < 0.001$). Moreover, the difference between the experimental and control groups was not significant in the hypothesis identification subscale at the pre-test level ($p = 0.392$), while the difference between the two groups was statistically significant at post-test level ($p < 0.001$).

Table 5.21: Pairwise Comparison Across Time for Both Groups in Hypothesis Identification

	Pre-test			Post-test			Within group(time)
	Mean	SD	Between group	Mean	SD	Between group	
Group			P value			P value	P value
Control	2.02	0.878	0.392	1.98	0.896	0.001	0.833
Experimental	1.89	0.791		2.52	0.825		<0.001

*The mean difference is significant at the .05 level Adjustment for multiple comparisons: Bonferroni.

Figure 5.5 contains two lines: the red line, which denotes the experimental group, and the blue line, representing the control group. The figure demonstrates a dramatic increase in the red line at the end of the learning intervention, while the blue line remained relatively flat by the end of the learning activity.

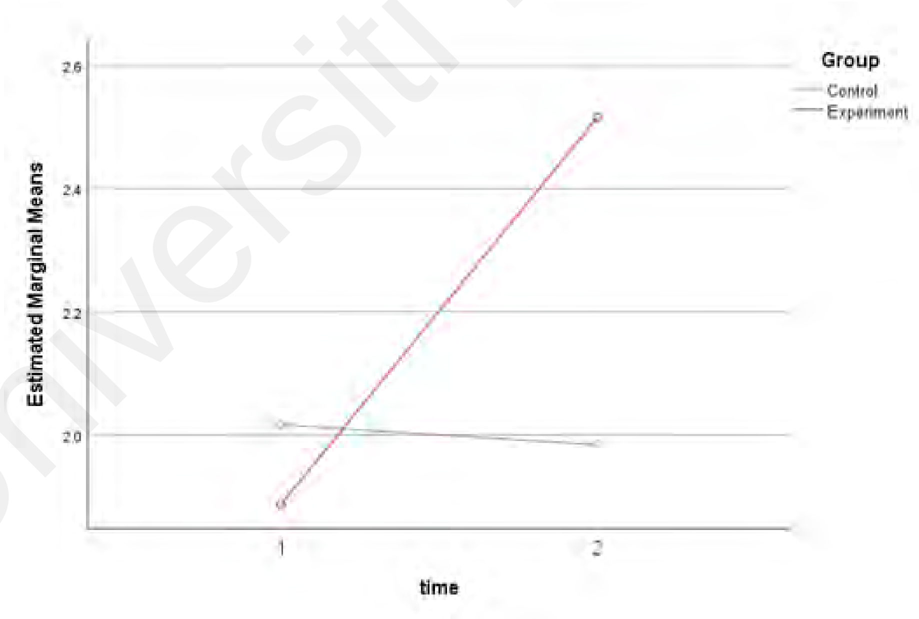

Figure 18.5: Means of Hypothesis Identification for Both Groups Across Time

5.4.1.2 Induction

The findings illustrated in Table 5.7 show the descriptive statistics (*M* and *SD*) of the induction subscale for pre- and post-test scenarios. The pre- and post-test scores of the experimental group students in the induction subscale were 2.77 (*SD* = 0.876) for the pre-test and 3.29 (*SD* = 0.818) for the post-test, while the mean scores of the control group were 2.53 (*SD* = 1.036) for the pre-test and 2.26 (*SD* = 1.007) for the post-test.

Table 5.22: Descriptive Statistics of Induction in Both Groups

Test	Group	N	Mean	SD
Pre test	Control	62	2.53	1.036
	Experimental	62	2.77	0.876
Post test	Control	62	2.26	1.007
	Experimental	62	3.29	0.818

As Table 5.8 demonstrates, the findings for the within-subjects effect (time) of the repeated-measures ANOVA were not significant ($F_{(1, 122)} = 0.947$, $p = 0.332$, $\eta^2 = 0.008$). The main effect of group was statistically significant ($F_{(1, 122)} = 31.326$, $p < 0.001$, $\eta^2 = 0.204$). In addition, the results suggested that the interaction between group and time was statistically significant ($F_{(1, 122)} = 10.103$, $p = 0.002$, $\eta^2 = 0.076$). These findings indicate that the change in induction scores was significantly different across the times.

Table 5.23: Summary of Repeated-Measures ANOVA for Induction Subscale

Source of variation	Df	MS	F	P Value	η2
Time	1	0.907	0.947	0.332	0.008
Group	1	25.165	31.326	0.000	0.204
Time * Group	1	9.681	10.103	0.002	0.076

A post-hoc Bonferroni test was performed to investigate this difference in students' scores by comparing the mean scores. According to the Bonferroni test results, displayed in Table 5.9, the difference of the induction mean scores between the pre-test and post-test in the control group was not statistically different ($p = 0.121$), while this difference in the experimental group reached the point of statistical significance ($p = 0.004$). Further, in the induction subscale, the difference between the experimental and control groups was not significant at the pre-test level ($p = 0.163$), while the difference between the two groups was statistically significant at the post-test level ($p < 0.001$).

Table 5.24: Pairwise Comparison Across Time for Both Groups in Induction Subscale

	Pre-test			Post-test			Within group(time)
	Mean	SD	Between group	Mean	SD	Between group	
Group			P value			P value	P value
Control	2.53	1.036	0.163	2.26	1.007	<0.001	0.121
Experimental	2.77	0.876		3.29	0.818		0.004

*The mean difference is significant at the .05 level Adjustment for multiple comparisons: Bonferroni.

Figure 5.6 shows a notable increase in the means of the experimental group across time; in contrast, the means of students in the control group experienced a considerable drop in performance at the end of the learning activity.

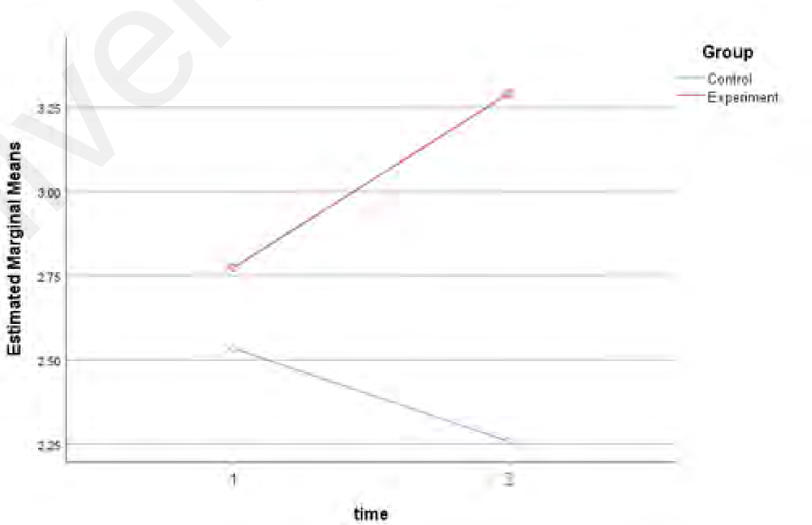

Figure 5.19: Means of Induction for Both Groups Across Time

5.4.1.3 Deduction

The results illustrated in Table 5.10 show the descriptive statistics (*M* and *SD*) of the deduction subscale for both pre- and post-tests. The pre- and post-test scores of the experimental group students in the deduction subscale were 2.53 (*SD* = 0.863) for the pre-test and 3.34 (*SD* = 0.700) for the post-test. The mean scores of the control group were 2.23 (*SD* = 0.895) for the pre-test and 2.29 (*SD* = 2.29) for the post-test.

Table 5.25: Descriptive Statistics of Deduction in Both Groups

Test	Group	N	Mean	SD
Pre test	Control	62	2.23	.895
	Experimental	62	2.53	.863
Post test	Control	62	2.29	1.136
	Experimental	62	3.34	.700

As Table 5.11 shows, the outcomes for the within-subjects effect (time) of the repeated-measures ANOVA was significant ($F_{(1, 122)} = 14.681, p < 0.001, \eta^2 = 0.107$). The main effect of group was statistically significant ($F_{(1, 122)} = 32.987$, p = < 0.001, η^2 = 0.213). These findings suggest that the interaction between group and time was statistically significant ($F_{(1, 122)} = 10.653$, $p = 0.001$, $\eta^2 = 0.08$), implying that the difference in deduction scores between the pre- and post-tests was statistically significant across time.

Table 5.26: Summary of Repeated-Measures ANOVA for Deduction Subscale

Source of variation	Df	MS	F	P Value	η2
Time	1	11.758	14.681	<0.001	0.107
Group	1	28.452	32.987	<0.001	0.213
Time * Group	1	8.532	10.653	0.001	0.08

To examine this difference, the researcher performed a post-hoc Bonferroni test to compare the mean scores. Based on the results shown in Table 5.12, the difference of the deduction mean scores between pre-test and post-test in the control group was not statistically different ($p = 0.689$), while the difference was statistically significant in the experimental group ($p < 0.001$). It should be noted that the difference between the experimental and control groups was not significant in the deduction subscale at the pre-test level ($p = 0.055$), while the difference between both groups was statistically significant at the post-test level ($p < 0.001$).

Table 5.27: Pairwise Comparison Across Time for Both Groups in Deduction

	Pre-test			**Post-test**			**Within group(time)**
	Mean	**SD**	**Between group**	**Mean**	**SD**	**Between group**	
Group			**P value**			**P value**	**P value**
Control	2.23	.895	0.055	2.29	1.136	<0.001	0.689
Experimental	2.53	.863		3.34	.700		<0.001

*The mean difference is significant at the .05 level. Adjustment for multiple comparisons: Bonferroni.

Figure 5.7 shows that both the experimental and control groups experienced an increase in their performance by the end of the learning intervention. However, it was the experimental group that achieved a significant increase in performance across time.

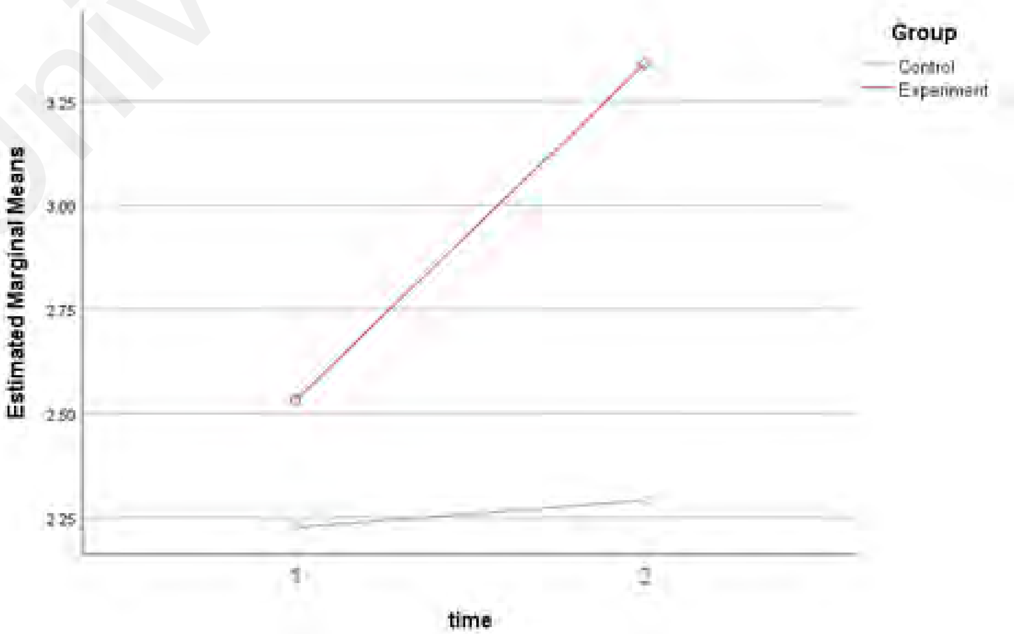

Figure 5.20: Means of Deduction for Both Groups Across Time

5.4.1.4 Explanation

The outcomes presented in Table 5.13 show the descriptive statistics (*M* and *SD*) of the evaluation subscale in pre- and post-tests. The pre- and post-test scores of the experimental group students in the evaluation subscale were 2.74 (SD = 0.867) and 3.13 (SD = 0.640), respectively, while the mean scores of the control group were 2.95 (SD = 0.876) at the pre-test and 2.40 (SD = 1.063) at the post-test levels.

Table 5.28: Descriptive Statistics of Explanation in Both Groups

Test	Group	N	Mean	SD
Pre test	Control	62	2.95	.876
	Experimental	62	2.74	.867
Post test	Control	62	2.40	1.063
	Experimental	62	3.13	.640

As Table 5.14 demonstrates, the findings for the within-subjects effect (time) of the repeated-measures ANOVA were not significant ($F_{(1, 122)} = 0.523$, $p = 0.471$, $\eta^2 = 0.004$). With regard to the main effect of group, the results indicated a statistical significance ($F_{(1,122)} = 17.599$, $p < 0.001$, $\eta^2 = 0.126$). These findings also suggest that the interaction between group and time was significant from a statistical perspective ($F_{(1, 122)} = 2573.987$, $p < 0.001$, $\eta^2 = 0.955$).

Table 5.29: Summary of Repeated-Measures ANOVA for Explanation Subscale

Source of variation	Df	MS	F	P Value	η2
Time	1	0.403	0.523	0.471	0.004
Group	1	13.565	17.599	0.000	0.126
Time * Group	1	1953.29	2573.987	0.000	0.955

A post-hoc Bonferroni test was implemented to further investigate this difference by comparing the mean scores. As shown in Table 5.15, the difference in the explanation mean scores between pre-test and post-test in the control group was statistically different ($p = 0.001$). However, this change in the mean scores was not in the desired direction, as it decreased from pre-test to post-test. Conversely, the difference achieved by the experimental group was statistically significant and in the desired direction ($p < 0.016$). With regards to this subscale, the difference between the experimental and control groups was not significant at the pre-test level ($p = 0.183$), while the difference between both groups was statistically significant at post-test level ($p < 0.001$).

Table 5.30: Pairwise Comparison Across Time for Both Groups in Explanation

	Pre-test			Post-test			Within group(time)
	Mean	SD	Between group	Mean	SD	Between group	
Group			**P value**			**P value**	**P value**
Control	2.95	.876	0.183	2.40	1.063	<0.001	0.001
Experimental	2.74	.867		3.13	.640		0.016

*The mean difference is significant at the .05 level. Adjustment for multiple comparisons: Bonferroni.

The performance of students in the experimental group increased over time. In contrast, the performance of students in the control group experienced a substantial decrease across time, as shown in Figure 5.8.

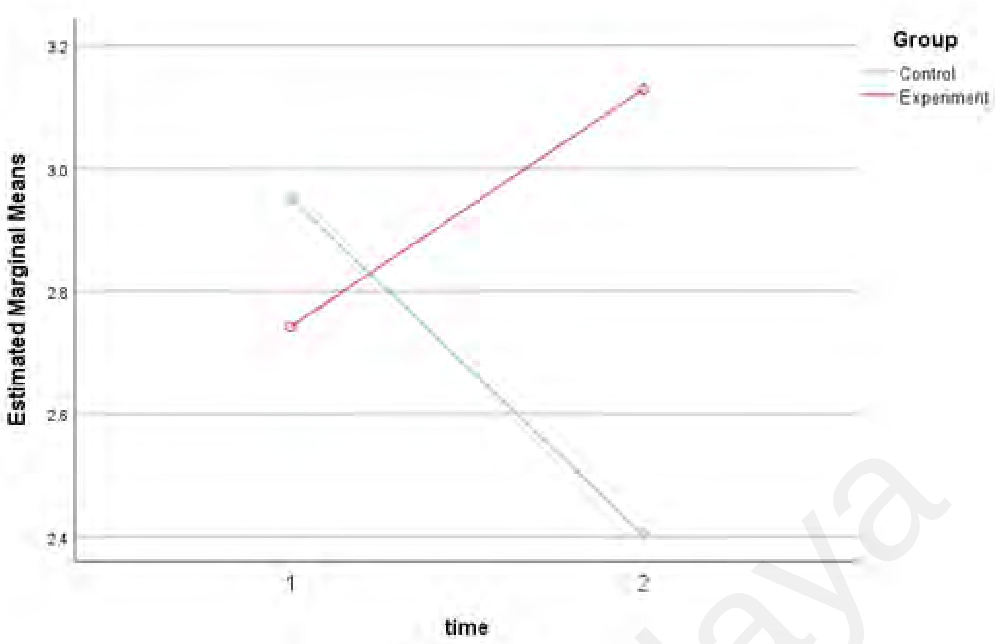

Figure 5.21: Means of Explanation for Both Groups Across Time

5.4.1.5 Evaluation

The outcomes in Table 5.16 denote the descriptive statistics (*M* and *SD*) of the explanation subscale in the pre- and post-tests. The pre- and post-test scores of the experimental group students in the explanation subscale were 2.52 (*SD* = 0.741) and 2.82 (*SD* = 0.758), respectively, while the mean scores of the control group were 2.26 (*SD* = 0.828) for the pre-test and 2.21 (*SD* = 1.026) for the post-test.

Table 5.31: Descriptive Statistics of Evaluation in Both Groups

Test	Group	N	Mean	SD
Pre test	Control	62	2.26	.828
	Experimental	62	2.52	.741
Post test	Control	62	2.21	1.026
	Experimental	62	2.82	.758

As shown in Table 5.17, the repeated-measures ANOVA results concerning the within-subjects effect (time) were not significant ($F_{(1, 122)} = 1.431$, $p = 0.234$, $\eta^2 = 0.012$). In relation to the main effect of group, the outcomes showed a statistical significance

($F_{(1,122)}$ = 16.553, p < 0.001, η^2 = 0.119). These findings suggest that the interaction between group and time was not statistically significant ($F_{(1,122)}$ = 2.705, p = 0.103, η^2 = 0.022).

Table 5.32: Summary of Repeated-Measures ANOVA for Evaluation Subscale

Source of variation	Df	MS	F	P Value	η2
Time	1	1.032	1.431	0.234	0.012
Group	1	11.758	16.553	0.000	0.119
Time * Group	1	1.952	2.705	0.103	0.022

Although the interaction between group and time was not statistically significant, this study nonetheless performed the post-hoc test (Bonferroni). The findings, as shown in Table 5.18, indicate the mean scores between pre-test and post-test in the control group were not statistically different (p = 0.752), while there was a significant difference in the experimental group (p = 0.047). Further, in the evaluation subscale, the difference between the experimental and control groups was not significant at the pre-test level (p = 0.070), while the difference between both groups was statistically significant at post-test level (p < 0.001).

Table 5.33: Pairwise Comparison Across Time for Both Groups in Evaluation

	Pre-test			Post-test			Within group(time)
	Mean	SD	Between group	Mean	SD	Between group	
Group			**P value**			**P value**	**P value**
Control	2.26	.828	0.070	2.21	1.026	<0.001	0.752
Experimental	2.52	.741		2.82	.758		0.047

*The mean difference is significant at the .05 level. Adjustment for multiple comparisons: Bonferroni.

With regard to the evaluation subscale, the mean of students who received the instruction via the DGBL method significantly increased across time, while students in the control group, who learned via the traditional method, did not experience such an increase; see Figure 5.9.

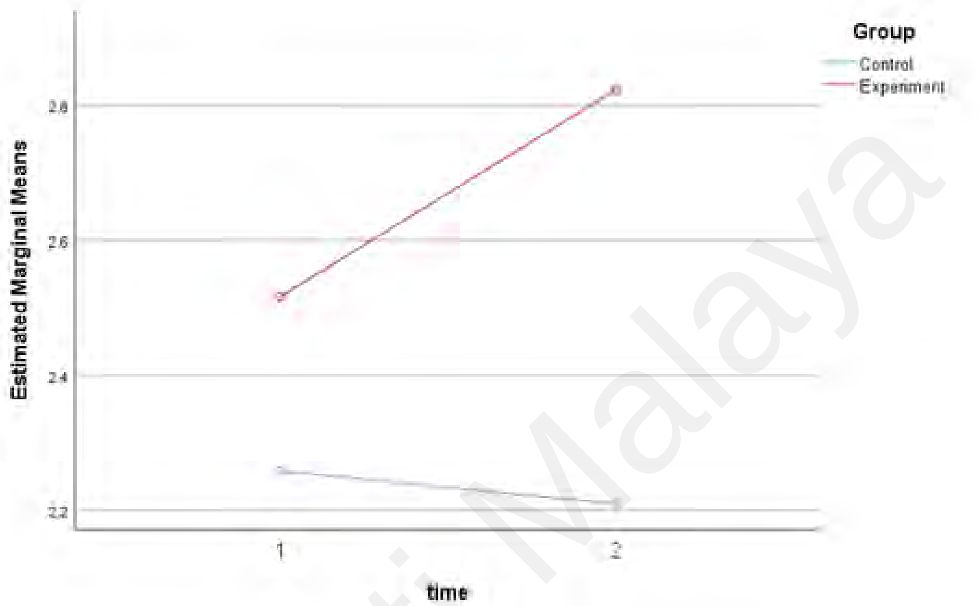

Figure 5.22: Means of Evaluation for Both Groups Across Time

5.4.1.6 Overall Critical Thinking Test

The outcomes in Table 5.19 present the descriptive statistics (*M* and *SD*) of the overall critical thinking test in the pre- and post-test scenarios. The pre- and post-test scores of the experimental group students were 12.45 (*SD* = 2.208) and 15.10 (*SD* = 2.125), respectively, while the control group's mean scores were 11.98 (*SD* = 1.877) for the pre-test and 11.15 (*SD* = 3.007) for the post-test.

Table 5.34: Descriptive Statistics of the Overall Critical Thinking Test in Both Groups

Test	Group	N	Mean	SD
Pre test	Control	62	11.98	1.877
	Experimental	62	12.45	2.208
Post test	Control	62	11.15	3.007

	Experimental	62	15.10	2.125

As shown in Table 5.20, the repeated-measures ANOVA results for the within-subjects effect (time) were statistically significant ($F_{(1, 122)} = 9.289, p = 0.003, \eta^2 = 0.071$). Regarding the main effect of group, the results revealed a statistical significance ($F_{(1, 122)} = 54.702, p < 0.001, \eta^2 = 0.31$). Furthermore, these findings indicate that the interaction between group and time was statistically significant ($F_{(1, 122)} = 34.551, p < 0.001, \eta^2 = 0.221$), suggesting that the differences of critical thinking scores between the pre- and post-tests reached the point of statistical significance.

Table 5.35: Summary of Repeated-Measures ANOVA for the Overall Critical Thinking Test in Both Groups

Source of variation	Df	MS	F	P Value	η2
Time	1	50.581	9.289	0.003	0.071
Group	1	302.726	54.702	0.000	0.31
Time * Group	1	188.129	34.551	0.000	0.221

To evaluate this difference, the researcher applied a post-hoc Bonferroni test to compare the mean scores of the critical thinking test. According to the results presented in Table 5.21, the difference in the control group students' critical thinking test between pre- and post-test was significant ($p < 0.05$). However, this difference does not indicate an improved performance: on the contrary, it reflects a decline in students after the learning activity. For the experimental group, meanwhile, the difference was statistically significant and in the desired direction ($p < 0.001$). Finally, with respect to the overall critical thinking assessment, the difference between the experimental and control groups was not significant at the pre-test level ($p = 0.206$), while the difference between both groups reached the point of statistical significance at post-test level ($p < 0.001$).

Table 5.36: Pairwise Comparison Across Time for Both Groups in the Overall Critical Thinking Test

	Pre-test			Post-test			Within group(time)
	Mean	SD	Between group	Mean	SD	Between group	
Group			P value			P value	P value
Control	11.98	1.877	0.206	11.15	3.007	<0.001	0.048
Experimental	12.45	2.208		15.10	2.125		<0.001

*The mean difference is significant at the .05 level. Adjustment for multiple comparisons: Bonferroni.

As shown in Figure 5.10, the overall mean scores of the critical thinking test imply that students in the experimental group witnessed a significant increase in their mean score across time. Conversely, there was a significant decrease in the overall mean score of students in the control group. Consequently, based on these findings, the null hypothesis is rejected, and the alternative hypothesis is supported: the proposed DGBL approach significantly improves students' critical thinking.

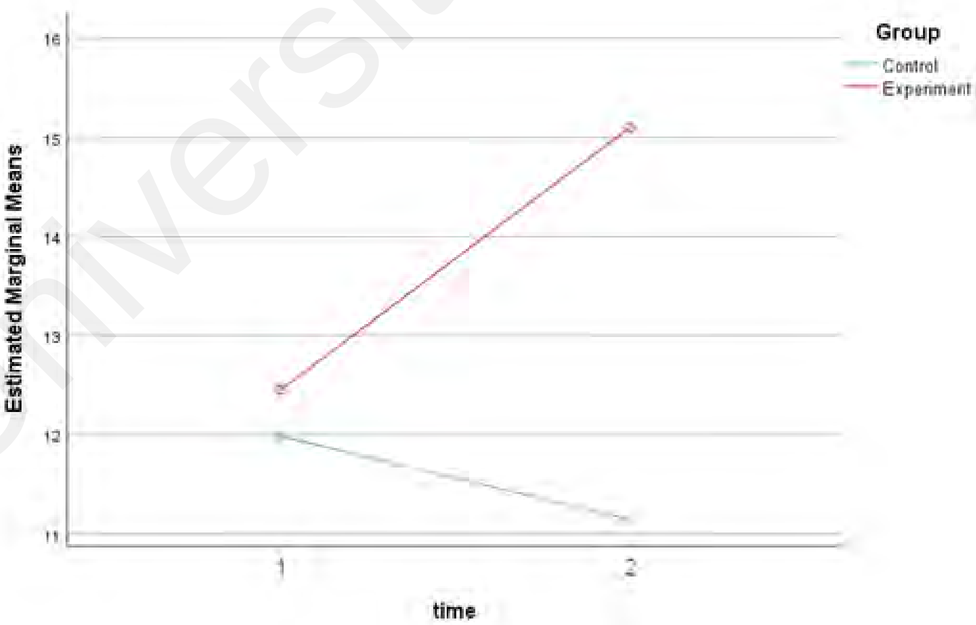

Figure 5.23: Means of Overall Critical Thinking for Both Groups Across Time

5.4.2 Science Subject Test

The following sections examine the following research hypotheses:

$H2_0$: The gaming application which is developed based on the proposed DGBL framework does not enhance students' knowledge acquisition.

H_{2a}: The gaming application which is developed based on the proposed DGBL framework does enhance students' knowledge acquisition.

5.4.2.1 Science Subject Test Analysis

Table 5.22 highlights the descriptive statistics (*M* and *SD*) of the science subject test for pre- and post-tests. The pre- and post-test scores of the experimental group were 56.66 (*SD* = 11.584) at the pre-test level and 77.50 (*SD* = 10.835) at the post-test level. The mean scores of the control group were 55.32 (*SD* = 11.270) for the pre-test and 57.68 (*SD* = 10.820) for the post-test.

Table 5.37: Descriptive Statistics of the Science Subject Test in Both Groups

Test	Group	N	Mean	SD
Pre test	Control	62	55.32	11.270
	Experimental	62	56.66	11.584
Post test	Control	62	57.68	10.820
	Experimental	62	77.50	10.835

The findings shown in Table 5.23 suggest that the repeated-measures ANOVA outcome for within-subjects effect (time) was significant ($F_{(1, 122)} = 9.289$, $p = 0.003$, $\eta^2 = 0.071$). Similarly, the main effect of group was statistically significant ($F_{(1, 122)} = 54.702$, $p < 0.001$, $\eta^2 = 0.31$). These findings indicate that the interaction between group and time was statistically significant ($F_{(1, 122)} = 34.551$, $p < 0.001$, $\eta^2 = 0.221$); thus, the change of science subject scores in both groups was significantly different across the times (i.e., pre- and post-test).

Table 5.38: Summary of Repeated-Measures ANOVA for the Science Subject Test in Both Groups

Source of variation	Df	MS	F	P Value	η2
Time	1	8338.081	159.211	0	0.566
Group	1	6940.903	35.509	0	0.225
Time * Group	1	5295.629	101.117	0	0.453

The post-hoc test (Bonferroni) was conducted to shed more light on this difference by comparing the mean values of the science subject scores. The difference of the science subject mean score between pre-test and post-test in the control group was not statistically different (p = 0.072). In contrast, this difference was statistically significant in the experimental group (p < 0.001). In relation to the science subject scores, the difference between students in the experimental and students control groups was not significant at the pre-test level (p = 0.515), while the difference between both groups was statistically significant at post-test level (p < 0.001); see Table 5.24.

Table 5.39: Pairwise Comparison Across Time for Both Groups in the Science Subject Test

	Pre-test			Post-test			Within group(time)
	Mean	SD	Between group	Mean	SD	Between group	
Group			P value			P value	P value
Control	55.32	11.270	0.515	57.68	10.820	<0.001	0.072
Experimental	56.66	11.584		77.50	10.835		<0.001

* The mean difference is significant at the .05 level Adjustment for multiple comparisons: Bonferroni.

Finally, with respect to students' achievements in the science subject test, Figure 5.11 shows that students in both groups made progress across time. However, only students in the experimental groups who learned via the DGBL approach achieved significant gains across time. Indeed, the achievements of students in the control group did not reach the point of statistical significance. These findings are consistent with the second alternative hypothesis. Hence, the second null hypothesis is rejected, as only the students in the experimental group made significant enhancements in their science subject scores.

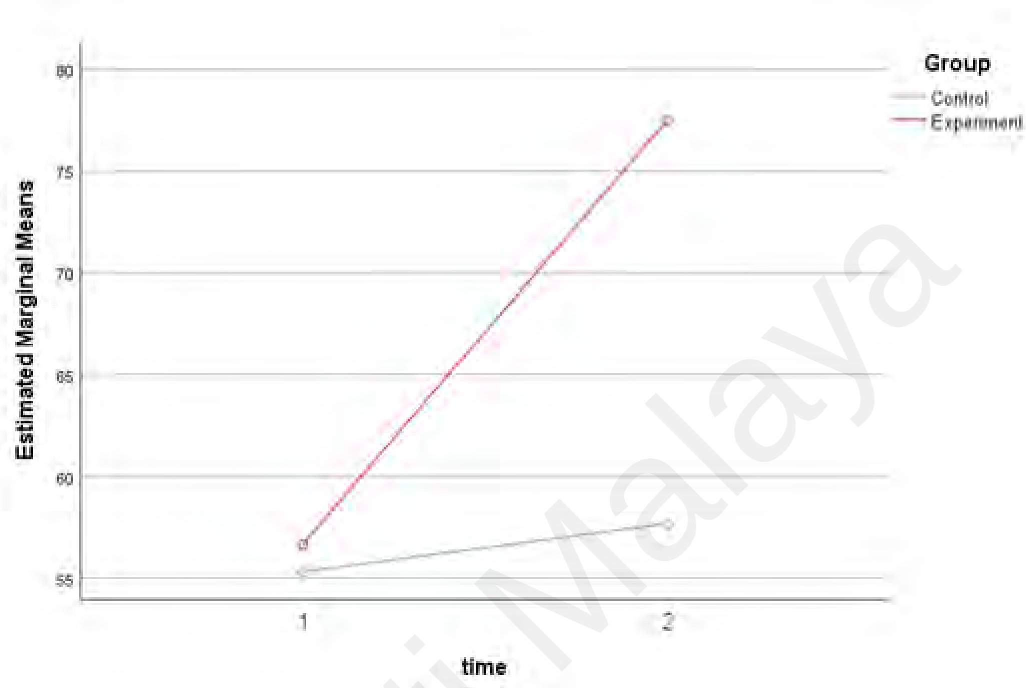

Figure 5.24: Means of Science Subject Scores for Both Groups Across Time

5.5 Pearson Correlation Analysis

The Pearson correlation is a statistical measure that examines the strength and direction of relationships between two variables (Schober & Schwarte, 2018). In this approach, the correlation coefficient can range from -1 to +1, where a value of -1 suggests a perfect negative correlation, a value of +1 suggests a perfect positive correlation, and a value of 0 indicates no correlation at all. This study utilised Guilford and Fruchter's (1973) criteria to determine the strengths of associations between variables, as seen in Table 5.25.

Table 5.40: Criteria for Interpreting Strength of Relationship Between Two Variables

R	Strength of Relationship
<.20	Slight relationship

.20-.40	Low correlation, definite but small
.40-.70	Moderate correlation, substantial relationship
.70-.90	High correlation, marked relationship
>.90	Very high correlation, very dependable relationship

As Table 5.26 shows, there were no significant correlations between the science subject test and the critical thinking test or any of its subscales at the pre-test level. However, by the end of the learning activity, there were positive and significant correlations between the science subject test and the critical thinking test ($r = 0.315$, $p < 0.01$), the induction subscale ($r = 0.376$, $p < 0.01$), and the deduction subscale ($r = 0.293$, $p < 0.01$). The values of these significant correlations range between 0.293 and 0.376. According to Guilford and Fruchter's (1973) model, the values indicate a small but definitive correlation between science subject and critical thinking.

Table 5.41: Science Subject Test with Critical Thinking and Critical Thinking Subscales

Variable	R value (pre-test)	Sig. (2-tailed)	R value (post-test)	Sig. (2-tailed)
Hypothesis identification subscale	0.074	0.412	0.031	0.735
Induction subscale	-0.071	0.436	0.376**	0.000
Deduction subscale	-0.022	0.807	0.293**	0.001
Explanation subscale	0.08	0.378	0.133	0.139
Evaluation subscale	-0.16	0.077	0.173	0.055
Overall critical thinking test	-0.04	0.658	0.315**	0.000

* Correlation is significant at the 0.05 level (2-tailed), ** Correlation is significant at the 0.01 level (2-tailed)

5.6 Summary

This research study developed a DGBL framework and evaluated its effectiveness in a quasi-experimental study conducted in a public elementary school in Malaysia. The results show that the data obtained from participants were normally distributed and students from both groups had equivalent critical thinking abilities and science knowledge before the start of the intervention. Meanwhile, the results from repeated-measures ANOVA tests indicated that the students in the experimental group significantly outperformed their peers in the control group in terms of critical thinking and science subject knowledge. In other words, after the learning activity, students who were guided by the proposed DGBL approach performed considerably better than students who were led by the traditional teaching method.

CHAPTER 6: DISCUSSION AND CONCLUSION

6.1 Introduction

These research findings could go a long way to assist educators, researchers, and developers in designing and implementing an effective DGBL application that improves students' critical thinking skills in the subject of primary-level science. This chapter presents a conclusion of the findings and offers some recommendations towards yielding more illuminative results in future works.

6.2 Discussion of Results

This part of the study answers the following research question 'How to evaluate the DGBL application in order to validate the proposed framework?'. Notwithstanding the literature's frequent observations that DGBL could promote students' academic achievements and has the potential to improve students' 21st-century skills such as critical thinking, the present study seeks to confirm these suggestions to determine whether the excitement is justified.

The proposed framework should be validated in an actual learning environment to answer the abovementioned research question effectively. Therefore, the researcher applied a quasi-experiment design that targeted Malaysian fifth-grade students. At the end of the implementation phase, the evaluation phase began, which comprised two sections. Firstly, prior to data analysis, numerous normality tests were conducted. Findings from these tests indicated that the data were normally distributed. Secondly, the homogeneity test results revealed no significant differences between students from both groups concerning their overall critical thinking test, the critical thinking test subscales, and the science subject test. Subsequently, students' pre- and post-test scores were analysed using the repeated-measures ANOVA and Bonferroni tests, which revealed significant differences and indicated that students who learned via the gaming application significantly outperformed their peers, who received their instructions via the traditional method, in the overall critical thinking test, every subscale of the critical thinking test, and the science subject test.

These results were consistent with the findings from previously published studies by Hwang and Chen (2017) and Yang and Chang (2013), who recognised gaming as an activity that could improve students' critical thinking abilities. Moreover, the significant improvement shown by students in the experimental group concerning their post-test score in the science subject test is corroborated by the findings of previous studies. These existing research works reported that educational computer games could improve students' acquisition of curricular science knowledge (Chu & Chang, 2013; Sung et al., 2015; Sung & Hwang, 2013, 2018).

The results of this study – especially with respect to students in the experimental group – show that students reacted positively to Ecoship Endeavour, as the critical thinking scores of experimental group students significantly advanced (pre-test = 12.45, post-test = 15.10, $p < 0.001$). In a similar manner, the scores of experimental group students in the science subject test improved significantly (pre-test = 56.66, post-test = 77.50, $p < 0.001$). The majority of students stated that Ecoship Endeavour was effective and entertaining. A possible explanation for this reaction could be attributed to the game design and the components embedded within the game. Although the game focused primarily on critical thinking, Ecoship Endeavour included other engaging activities, such as interactions with other teams, score comparisons, and attempts to obtain better and higher scores than other teams.

In addition, the performance of students in the experimental group corresponds to what has been reported by Malaysian students in the TIMSS assessments. In the last three cycles of TIMSS, Malaysian students demonstrated a very promising attitude towards science learning which was above the international average of 35%, as the percentage of students who reported that they "Very much like learning science" was 42% in 2011, 51% in 2015, and 46% in 2019 (Phang et al., 2020). This suggests that the traditional method has been ineffective in cultivating students' willingness to learn, which consequently hindered the development of their critical thinking. Hence, providing students with a DGBL application that is fun to play and rooted in specific learning principles could capitalise on students' positive attitude, hence, improving their critical thinking and science subject knowledge.

The results also indicated that the critical thinking means of students in the control group experienced a significant decline: the pre-test score of students in the control group at the beginning of the intervention was 11.98, and by the end of the learning intervention, their post-test score declined to 11.15 ($p = 0.048$). Similarly, the performance of control group students in the science subject test did not significantly progress after the learning activity (pre-test = 55.32, post-test = 57.68, $p = 0.072$).

This observation regarding the performance of students in the control group can be linked to the traditional method of teaching. This method emphasises memorisation and preparing students for tests. The findings of the present study are in line with numerous others that have criticised this method of instruction (Li & Tsai, 2013). For example, Bikić, Maričić, and Pikula (2016) and Voskoglou and Salem (2020) noted that students' exposure to complex problems is extremely limited. Thus, using the traditional method does not advance the development of students' problem-solving skills, conceptual understanding, or critical analysis competencies. In addition, the findings of this study are also in line with what has been achieved in several international large-scale assessments, which revealed that the performance of Malaysian students was poor and uninspiring not only in the subject of science but also in other subjects (e.g., mathematics and reading). Although the performance of students in the control group is unsatisfactory, it was not surprising, as Seman, Yusoff, and Embong (2017) noted that teaching and learning low-level thinking skills still dominates the Malaysian schooling system. In fact, Sardareh and Saad (2013) stated that even when Malaysian teachers utilise open-ended questions, these questions targetted lower cognitive skills and they expected students to provide specific and precise answers. Such practices, according to Sardareh and Saad (2013) do not help students to use or improve their thinking skills. Therefore, when Mullis et al. (2012) analysed Malaysian students' performance in the TIMSS they revealed that only 2 – 10% of students can interpret information and draw generalisations when solving complex problems. Moreover, Tan, Ismail, and Abidin (2018) noted that while the Malaysian curricula are incorporating learning tasks that focus on knowing, applying, and reasoning, the majority of these tasks focus on the knowing domain which requires less mastery of higher order thinking skills compared to the other domain of applying and reasoning. As such, Tan et al. (2018) argued that the poor performance of Malaysians students in learning tasks that demand higher order thinking skills should not be a surprise for teachers and educators.

In addition to the outcomes of the cognitive domains in the TIMSS assessment, the findings of this study correspond to what have been observed by a number of critical thinking studies. For example, in an intervention that lasted for eight weeks, Darby and Rashid (2017) noted that when students were taught via the traditional method of teaching their critical thinking dropped from (M = 51.66, SD = 3.05) in the pre-test, to (M = 50.80, SD = 2.61) in the post-test. Additionally, according to Kanbay and Okanlı (2017), the traditional method does not develop or advance students' critical thinking, as the findings of their learning intervention demonstrated a significant decrease in students critical thinking from (M = 260.79, SD = 3.76) on the pre-test, to (M = 258.18, SD = 3.83) on the post-test. In a similar manner, when the traditional method was implemented for eight weeks, Arsal (2017) recorded a sharp decline in students' critical thinking from (*M* = 219.04, *SD* = 17.43) at the beginning of the learning activity, to (*M* = 216.79, *SD* = 21.62) at the end of the learning activity.

Further, the findings of the current study are also supported by what has been documented by a number of studies that focused on other 21st-century skills. For instance, in a learning intervention that lasted for three weeks, Yoo and Park (2015) reported that students' motivation to learn decreased significantly from (*M* = 99.53, *SD* = 10.85) in the pre-test, to (*M* = 98.29, *SD* = 12. 97) in the post-test, when they engaged in higher order thinking exercises such as problem-solving that was administered via the traditional method. In addition, Kanbay and Okanlı (2017) also noted that the problem-solving skills of students who were taught via the traditional method of teaching decreased from (*M* = 86.31, *SD* = 3.15) in the pre-test, to (*M* = 87.71, *SD* = 2.83) in the post-test. However, it should be noted that problem-solving skills of students fall as their scores increase, as the utilised assessment was marked in reversed manner. Furthermore, in a recent study, Lee, Lin, Hwang, Fu, and Tseng (2021) documented a significant decline in students creativity when they were guided by the traditional method of teaching from (*M* = 90.70, *SD* = 1.97) in the pre-test, to (*M* = 89.58, *SD* = 3.01) in the post-test, over a period of three weeks.

In other words, the outcomes from a number of academic studies and large scale international assessments suggest that when students are taught via the traditional method, they will not be afforded any meaningful chances to practice their critical thinking and teachers also will not be able to teach or help students improve their critical thinking. As such, students' critical thinking and other higher-order thinking skills will inevitably drop as students progress in their education.

6.3 Contributions of the Study

In essence, the main contribution of this thesis to the body of knowledge stems from examining the empirical effectiveness of the proposed ICMDCR framework on students' critical thinking. Accordingly, the output of this research can be summarised as follows.

- It contributes to and extends the existing literature on DGBL in the context of critical thinking, a domain that remains largely unexplored.
- It offers empirical evidence regarding the effectiveness of the DGBL method in enhancing students' critical thinking skills in their early stages of education.
- This research study has successfully designed and developed a DGBL application using the proposed framework.
- The study offers the development of a framework to guide the process of designing an effective DGBL application to learn primary-level science.
- The study findings may be used to assist primary school teachers and educators in applying a DGBL application to teach science to young learners in an enjoyable and interactive way.

6.4 Limitations

While this study has presented helpful and relevant findings, it was subject to several limitations. Limitations are weaknesses or potential issues with a study that have been identified by the researcher (Creswell, 2012).

This study has the following limitations.

- In this study, all participants were students from the same school. The use of one school could limit the generalisation of results.

- The research sample lacked male participants, as the school where the study was based was an all-girls school. This absence of male participants could also limit the generalisation of the results.
- A number of schools were reluctant to participate in this study, as many teachers are unable to use computers effectively in their teaching practices. In addition, some teachers were unable to join this study due to their hectic schedule and the numerous tasks associated with the traditional method of teaching.
- The learning intervention in this study lasted for three weeks. Had this activity continued for a longer duration of time, the proposed game could have included more curricular science topics, which would have produced more conclusive results.
- Furthermore, the non-randomisation of the research sample was another limitation. In this study, students were divided into the experimental and control group on a class per class basis. As such, further studies are needed to employ a more rigours research design such as RCTs to ensure the randomisation of the research sample.

6.5 Implications

Despite the abovementioned limitations, the findings of this study highlighted the promising potential of the DGBL approach. As such, this study yields the following implications.

Firstly, the research offers implications for both schools and teachers. Teachers could utilise DGBL applications to provide students with opportunities through which they can explore, experiment, inquire, and learn independently at their own pace and have the freedom to reexamine the learning material multiple times at their leisure by capitalising on the learner-centred aspect native to this method of instruction.

In addition to demonstrating the effectiveness of the DGBL approach, this study also presented a detailed clarification of how the components of the ICMDCR framework could facilitate and improve students' critical thinking. Based on this, researchers and game developers can employ the ICMDCR framework to design DGBL applications to improve students' critical thinking in other grades or other curricular subjects.

Further, this study empirically identified the active role of DGBL in teaching and learning. Therefore, leaders and those in decision making positions in the Malaysian educational institution need to pay closer attention to the DGBL method and recognise the key role this instructional approach plays in shaping students' 21st-century skills and preparing them for university education and future careers in competitive and global markets.

Secondly, the study has implications for game developers and educators. Since DGBL can improve and advance students' critical thinking skills, game developers who are interested in designing and developing DGBL applications should aim to situate students in rich playing experiences. To provide such experiences, designers should prioritise increasing students' engagement and interest in educational digital games in a similar way to successful and popular commercial video games. Indeed, it is essential for developers to pay close attention to the factors found to be effective in generating and improving students' interest and engagement in DGBL applications while also improving their 21st-century skills and contributing positively to their learning achievements.

Moreover, to maximise the efficiency of the DGBL approach, it is not enough to design gaming applications based on specific learning frameworks and engaging graphics, but it is also essential for educators to modify their teaching techniques in accordance with students' needs and provide them with sufficient and individualised support while they learn via the DGBL approach.

From a practical and realistic point of view, not every school can utilise the DGBL approach for a number of reasons such as time constraints, financial constraints, technical literacies, and organisational support. However, teachers and educators can implement a number of the activities provided by the ICMDCR framework in a traditional classroom environment. For example, teachers can encourage and teach students how to raise thoughts provoking questions, which is the essence of inquiry-based learning activities.

Interaction with teammates would help students engage in meaningful discussions and listen to multiple streams of thoughts, as a result, students will broaden their perspectives, expand their horizons, and improve their communication skills.

In the traditional method, students rarely receive feedback from their teachers, therefore, adopting such a practice where teachers provide guidance and help to students when they need it would produce a better and more dynamic learning experience.

In addition, when teachers interact with students, they are encouraged to avoid questions that require one definitive answer. By doing so, they would prompt students to explore several alternatives and judge their viabilities in order to determine which one is the most suitable path to follow.

When students become accustomed to such practices, teachers can provide students with more challenging questions and exercises. By doing so, students would cooperate among themselves to overcome these obstacles, hence, enhancing their problem-solving skills. Finally, teachers are highly encouraged to reward students when they fulfill their learning tasks successfully by verbally commending them or by providing them with other forms of recognition such as badges, when teachers embrace such techniques, students will feel a sense of achievement and that their efforts are acknowledged by their teachers.

6.6 Future Work

This study highlights one aspect of designing a DGBL application to improve and promote students' critical thinking skills. Nonetheless, more research is required to enhance the effectiveness of this method of instruction. As such, further empirical studies with a larger number of participants are necessary to provide further evidence that the proposed ICMDCR framework effectively enhances students' critical thinking abilities. Additionally, more empirical research works are warranted for a more extended period to offer researchers the ability to collect a larger corpus of data. Such data may provide additional evidence pertaining to the effectiveness of the ICMDCR framework in improving students' critical thinking competencies.

6.7 Recommendations

Several recommendations can be proposed based on the findings of this study. These recommendations aim to assist scholars and educators in conducting research more efficiently and in improving the state of DGBL integration in schools and educational institutes. The recommendations are listed below.

- Schools could capitalise on the popularity of video games among students and the recent ubiquitous employment of technology in education to develop training programmes for teachers. These programmes could improve teachers' ICT skills and e-learning knowledge, enhance their self-confidence, familiarise them with using DGBL, and communicate the potential advantages of implementing this method in an educational context.

- Establishing a functioning channel of communication with teachers, educators, and students is vital to successfully develop and conduct a DGBL intervention. Teachers and educators can help researchers to identify the educational value of the components intended to be incorporated in their proposed DGBL framework, while students may highlight the motivational and entertaining elements they would like to have in the game.
- Teaching students 21st-century skills, such as critical thinking, creativity, problem-solving, communication, and collaboration, would benefit them greatly, particularly within formal classroom environments and at the K–12 level. Encouraging students to contemplate, reflect, and analyse facilitates powerful mental tools that would foster the determination, purpose, and discipline necessary to persevere and achieve their goals. This can be achieved when Malaysian teachers significantly reduce their overreliance on the spoon-feeding approach and start embracing more appropriate teaching methods that appeal to students and improve their 21st-century skills such as critical thinking.
- As previously mentioned, it is a difficult task to develop a DGBL application based on existing DGBL frameworks, because most of these frameworks do not provide researchers or serious games developers with the necessary information and steps to conduct such task effectively. As a consequence, this study recommends that upcoming studies provide such information and guidance. This, in turn, could play a key role in advancing DGBL research and improving the effectiveness of DGBL applications.

6.8 Summary of Findings

Numerous studies have applied DGBL. Indeed, a review of the literature indicated that DGBL could improve students' learning performance and motivation, and students reported that they enjoyed using game-based learning. However, the application of DGBL within the domain of critical thinking is still nascent and in a developmental stage, despite few studies producing promising evidence concerning the potential of DGBL utilisation in improving students' critical thinking, especially, when the gaming application is appropriately designed and implemented. As such, in this research, the first step was to develop a suitable DGBL framework that could provide a roadmap for producing an effective design for a DGBL application.

The first research objective entailed developing a DGBL framework for critical thinking based on the findings and highlighted observations established by reviewing prior research studies related to DGBL interventions in general, and DGBL interventions in the domain of science in particular, as well as research works that developed DGBL frameworks. The first research question, meanwhile, involved determining the essential components needed when designing the framework. In response to this objective, this research developed the ICMDCR framework to improve students' critical thinking skills.

In this study, a DGBL application was designed and developed in accordance with the ICMDCR framework to improve students' critical thinking skills. The purpose was to address the second research objective, which involved designing and developing a DGBL application that embeds the ICMDCR framework. This phase also addressed the second research question: 'How to develop a DGBL application using the proposed framework?'.

This study employed repeated-measures ANOVA and post-hoc Bonferroni tests to investigate students' critical thinking skills and science subject scores. These tests served to address the third and final research objective, which was centred around evaluating the DGBL framework by testing the effectiveness of the application among primary school students. Moreover, these tests helped to answer the final research question, which addressed ‘How to evaluate the DGBL application in order to validate the proposed framework?’.

The critical thinking test provided an opportunity to gain insight into students’ abilities to learn and practice one of the most essential and sought-after 21st-century skills: critical thinking. In addition, the science subject test served to demonstrate how students would perform under the regular tests administered by their science teachers. The findings suggested that students who learned via the DGBL approach exhibited signs of significant improvement in the critical thinking skills assessment (pre-test = 12.45, post-test = 15.10, $p < 0.001$). In addition, students in the experimental group achieved significant progress in every subscale of the critical thinking test. Further, Experimental group students also exhibited significant gains in their science subject test (pre-test = 56.66, post-test = 77.50, $p < 0.001$). In contrast, the students in the control group demonstrated a significant decline in the critical thinking test (pre-test = 11.98, post-test = 11.15, $p = 0.048$). With regards to the subscales of the critical thinking test, students in the control group demonstrated no significant attainment in any subscale. Further, in the science subject test, students in the control group showed improvement in their scores, however, their progress did not reach the point of statistical significance (pre-test = 55.32, post-test = 57.68, $p > 0.05$).

REFERENCES

Abd Majid, M. Z., Hussin, M., Norman, M. H., & Kasavan, S. (2020). The employability skills among students of Public Higher Education Institution in Malaysia. *Geografia-Malaysian Journal of Society and Space*, *16*(1).

Abdul Jabbar, A. I., & Felicia, P. (2015). Gameplay engagement and learning in game-based learning: A systematic review. *Review of educational research*, *85*(4), 740-779.

Abdulai, I. A., & Shafiwu, A. B. (2014). Participatory decision making and employee productivity. A case study of community banks in the upper east region of Ghana. *Business and Economics Journal*, *5*(3), 1.

Abduljabbar, A. L. (2019). *The Predictive Utility of Critical Thinking Skills on Language Proficiency in Adult English as a Second Language Learners* (Doctoral dissertation, Grand Canyon University).

Abdullah, A., & Peters, R. F. (2015). Malaysia's Post-PISA2012 Direction. *International Journal of Culture and History*, *1*(1), 15-20.

Abdullah, R. N., Yaacob, R. A. I. R., Hashim, A. J. C. M., Hussain, I. A., & Roslan, A. (2019). Intensifying experiential learning with dynamic learning styles in traditional classroom for technical and vocational students. *Journal of Technical Education and Training*, *11*(4), 1-10.

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research*, *78*(4), 1102-1134.

Abt, C.C. (1971), *Serious Games*, Viking Press, New York, NY.

Advisory Council on Economic Growth, 2017. Learning Nation: Equipping Canada's Workforce with Skills for the Future. Retrieved from. https://www.budget.gc.ca/aceg-ccce/pdf/learning-nation-eng.pdf

Agdas, S. (2013). *Effects of problem based learning on development of critical thinking skills and dispositions in engineering*. University of Florida.

Ahmar Ahmad, S. (2019). Exploring Lecturers' Readiness for 21st Century Education in Malaysian Higher Learning Institutions. *European Journal of Teaching and Education*.

Ahmed, S. & Parsons, D. (2013). Abductive science inquiry using mobile devices in the classroom. *Computers & Education*, *63*, 62–72.

Ahrari, S., Samah, B. A., Hassan, M. S. H. B., Wahat, N. W. A., & Zaremohzzabieh, Z. (2016). Deepening critical thinking skills through civic engagement in Malaysian higher education. *Thinking Skills and Creativity*, *22*, 121-128.

Aleven, V., Myers, E., Easterday, M., & Ogan, A. (2010, April). Toward a framework for the analysis and design of educational games. In *2010 third IEEE international conference on digital game and intelligent toy enhanced learning* (pp. 69-76). IEEE.

Alexander, G., Eaton, I., & Egan, K. (2010). Cracking the code of electronic games: Some lessons for educators. *Teachers College Record*, *112*(7), 1830-1850.

Alkhafaji, A. (2018). *A critical analysis of mystery in videogames*. https://via.library.depaul.edu/cdm_etd/17/

Alyaz, Y., Spaniel-Weise, D., & Gursoy, E. (2017). A Study on Using Serious Games in Teaching German as a Foreign Language. *Journal of Education and Learning*, *6*(3), 250-264.

Ambigapathy, P., & Aniswal, A. G. (2005). University curriculum: An evaluation on preparing graduates for employment. *Higher Education Research Monograph*, *6*, 2005.

An, Y. (2018). The effects of an online professional development course on teachers' perceptions, attitudes, self-efficacy, and behavioral intentions regarding digital game-based learning. *Educational Technology Research and Development*, *66*(6), 1505-1527.

Anastopoulou, S., Sharples, M., Ainsworth, S., Crook, C., O'Malley, C., & Wright, M. (2012). Creating personal meaning through technology-supported science inquiry learning across formal and informal settings. *International Journal of Science Education*, *34*(2), 251-273.

Andreoli, R., Corolla, A., Faggiano, A., Malandrino, D., Pirozzi, D., Ranaldi, M., ... & Scarano, V. (2017). A framework to design, develop, and evaluate immersive and collaborative serious games in cultural heritage. *Journal on Computing and Cultural Heritage (JOCCH)*, *11*(1), 1-22.

Annetta, L. A. (2010). The "I's" have it: A framework for serious educational game design. *Review of General Psychology*, *14*(2), 105-113.

Apperley, T., & Walsh, C. (2012). What digital games and literacy have in common: a heuristic for understanding pupils' gaming literacy. *Literacy*, *46*(3), 115-122.

Arsal, Z. (2017). The impact of inquiry-based learning on the critical thinking dispositions of pre-service science teachers. *International Journal of Science Education*, *39*(10), 1326-1338.

Asbell-Clarke, J., Edwards, T., Rowe, E., Larsen, J., Sylvan, E., & Hewitt, J. (2012). Martian boneyards: scientific inquiry in an MMO game. *International Journal of Game-based Learning, 2(1)*, 52–76.

Badri, M., Alnuaimi, A., Mohaidat, J., Al Rashedi, A., Yang, G., & Al Mazroui, K. (2016). My science class and expected career choices—a structural equation model of determinants involving Abu Dhabi high school students. *International Journal of STEM Education*, *3*(1), 1-21.

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational technology research and development*, *53*(1), 86-107.

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, *70*, 65-79.

Behnamnia, N., Kamsin, A., & Ismail, M. A. B. (2020). The landscape of research on the use of digital game-based learning apps to nurture creativity among young children: A review. *Thinking Skills and Creativity*, *37*, 100666.

Behnamnia, N., Kamsin, A., Ismail, M. A. B., & Hayati, A. (2020). The effective components of creativity in digital game-based learning among young children: A case study. *Children and Youth Services Review*, *116*, 105227.

Bellanca, J. A. (Ed.). (2010). *21st century skills: Rethinking how students learn*. Bloomington, IN: Solution Tree Press.

Bikić, N., Maričić, S. M., & Pikula, M. (2016). The effects of differentiation of content in problem-solving in learning geometry in secondary school. *EURASIA Journal of Mathematics, Science and Technology Education*, *12*(11), 2783-2795.

Billieux, J., Van der Linden, M., Achab, S., Khazaal, Y., Paraskevopoulos, L., Zullino, D., & Thorens, G. (2013). Why do you play World of Warcraft? An in-depth exploration of self-reported motivations to play online and in-game behaviours in the virtual world of Azeroth. *Computers in Human Behavior*, *29*(1), 103-109.

Birk, M. V., Atkins, C., Bowey, J. T., & Mandryk, R. L. (2016, May). Fostering intrinsic motivation through avatar identification in digital games. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2982-2995).

BONNEY, C. N. R., & STERNBERG, R. R. J. (2016). Learning to think critically. In *Handbook of research on learning and instruction* (pp. 191-222). Routledge.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., ... & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, *94*, 178-192.

Branch, R. M. (2009). *Instructional design: The ADDIE approach* (Vol. 722). Springer Science & Business Media.

Breien, F. S., & Wasson, B. (2021). Narrative categorization in digital game-based learning: Engagement, motivation & learning. *British Journal of Educational Technology*, *52*(1), 91–111. https://doi.org/10.1111/bjet.13004

Bruce, B. & Casey, L. (2012). The Practice of Inquiry: A Pedagogical 'Sweet Spot' for Digital Literacy?. *Computers In The Schools*, *29*(1-2), 191–206.

Bugis, Y. M. (2018). Creating Digital Stories with Saudi Arabian Pre-Service Teachers: Using the Analysis, Design, Development, Implementation, and Evaluation Model to Promote Lesson Plan Development.

Bugis, Y. M. (2018). *Creating Digital Stories With Saudi Arabian Pre-Service Teachers: Using the Analysis, Design, Development, Implementation, and Evaluation Model To Promote Lesson Plan Development*. *May*, 186.

Bunchball, I. (2010). Gamification 101: An introduction to the use of game dynamics to influence behavior. *White paper*, *9*.

Butler, H. A. (2012). Halpern Critical Thinking Assessment predicts real-world outcomes of critical thinking. *Applied Cognitive Psychology*, *26*(5), 721-729.

Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century US workforce*. Partnership for 21st Century Skills. 1 Massachusetts Avenue NW Suite 700, Washington, DC 20001.

Chang, C. C., Warden, C. A., Liang, C., & Lin, G. Y. (2018). Effects of digital game-based learning on achievement, flow and overall cognitive load. *Australasian Journal of Educational Technology*, *34*(4).

Chang, C. Y., & Hwang, G. J. (2019). Trends in digital game-based learning in the mobile era: a systematic review of journal publications from 2007 to 2016. *International Journal of Mobile Learning and Organisation*, *13*(1), 68-90.

Chase, C. C., Malkiewich, L. J., Lee, A., Slater, S., Choi, A., & Xing, C. (2020). Can typical game features have unintended consequences? A study of players' learning and reactions to challenge and failure in an educational programming game. *British Journal of Educational Technology*, e13434.

Chen, C. H., Shih, C. C., & Law, V. (2020). The effects of competition in digital game-based learning (DGBL): a meta-analysis. *Educational Technology Research and Development*, *68*(4), 1855-1873.

Chen, C. M. (2021). A Searchable Spreadsheet for Educational Games in the Decision Sciences. *Decision Sciences Journal of Innovative Education*, *19*(3), 197-203.

Chen, L. X., & Sun, C. T. (2016). Self-regulation influence on game play flow state. *Computers in Human Behavior*, *54*, 341-350.

Chen, S. W., Yang, C. H., Huang, K. S., & Fu, S. L. (2019). Digital games for learning energy conservation: A study of impacts on motivation, attention, and learning outcomes. *Innovations in Education and Teaching International*, *56*(1), 66-76.

Chen, Y. C. (2017). Empirical study on the effect of digital game-based instruction on students' learning motivation and achievement. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(7), 3177-3187.

Cheng, M. T., & Annetta, L. (2012). Students' learning outcomes and learning experiences through playing a Serious Educational Game. *Journal of Biological Education*, *46*(4), 203-213.

Cheng, M. T., Chen, J. H., Chu, S. J., & Chen, S. Y. (2015). The use of serious games in science education: a review of selected empirical research from 2002 to 2013. *Journal of computers in education*, *2*(3), 353-375.

Cheng, M. T., Su, T., Huang, W. Y., & Chen, J. H. (2014). An educational game for learning human immunology: What do students learn and how do they perceive?. *British Journal of Educational Technology*, *45*(5), 820-833.

Chiang, T. H., Yang, S. J., & Hwang, G. J. (2014). Students' online interactive patterns in augmented reality-based inquiry activities. *Computers & Education*, *78*, 97-108.

Chorianopoulos, K., & Giannakos, M. (2014). Design principles for serious video games in mathematics education: from theory to practice. *International Journal of Serious Games, 1(3)*, 51–59.

Chou, T. L., Wu, J. J., & Tsai, C. C. (2019). Research Trends and Features of Critical Thinking Studies in E-Learning Environments: A Review. *Journal of Educational Computing Research*, *57*(4), 1038-1077.

Choy, S. C., & Cheah, P. K. (2009). Teacher perceptions of critical thinking among students and its influence on higher education. *International Journal of teaching and learning in Higher Education*, *20*(2), 198-206.

Choy, S. C., & Cheah, P. K. (2009). Teacher perceptions of critical thinking among students and its influence on higher education. *International Journal of teaching and learning in Higher Education*, *20*(2), 198-206.

Chu, H. C., & Chang, S. C. (2014). Developing an educational computer game for migratory bird identification based on a two-tier test approach. *Educational Technology Research and Development*, *62*(2), 147-161.

Cicchino, M. I. (2015). Using game-based learning to foster critical thinking in student discourse. *Interdisciplinary Journal of Problem-Based Learning*, *9*(2).

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of educational research*, *86*(1), 79-122.

Claro, M., Salinas, A., Cabello-Hutt, T., San Martín, E., Preiss, D. D., Valenzuela, S., & Jara, I. (2018). Teaching in a Digital Environment (TIDE): Defining and measuring teachers' capacity to develop students' digital information and communication skills. *Computers and Education*, *121*, 162–174.

Cleaves, A. (2005). The formation of science choices in secondary school. *International Journal of Science Education*, *27*(4), 471-486.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & education*, *59*(2), 661-686.

Cook, D. A., & Ellaway, R. H. (2015). Evaluating technology-enhanced learning: a comprehensive framework. *Medical teacher*, *37*(10), 961-970.

Cooksey, R. W. (2020). Illustrating Statistical Procedures: Finding Meaning in Quantitative Data. In *Illustrating Statistical Procedures: Finding Meaning in Quantitative Data*. https://doi.org/10.1007/978-981-15-2537-7

Cornish, E. (2004). *Futuring: The exploration of the future*. Bethesda, MD: World Future Society.

Corti, K. (2006). Games-based Learning; a serious business application. *Informe de PixelLearning*, *34*(6), 1-20.

Cota, É., Ribeiro, L., Bezerra, J. S., Costa, A., da Silva, R. E., & Cota, G. (2017). Using formal methods for content validation of medical procedure documents. *International Journal of Medical Informatics*, *104*, 10-25.

Create Games with Construct 2. (2019). Retrieved 2 December 2019, from https://www.scirra.com/

Crenshaw, P., Hale, E., & Harper, S. L. (2011). Producing intellectual labor in the classroom: The utilization of a critical thinking model to help students take command of their thinking. *Journal of College Teaching & Learning (TLC)*, *8*(7), 13-26.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston, MA: Pearson.

Crow, T. V. (2015). *A mobile game world for Māori language learning: a thesis presented in partial fulfilment of the requirements for the degree of Master of Information Sciences at Massey University, Albany Campus, Auckland, New Zealand* (Doctoral dissertation, Massey University).

Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience.* New York: Harper Perennial.

Csikszentmihalyi, M. (2014). Toward a psychology of optimal experience. In *Flow and the foundations of positive psychology* (pp. 209-226). Springer, Dordrecht.

Da Wan, C., Sirat, M., & Razak, D. A. (2018). Education in Malaysia towards a developed nation.

Daniels, M. E. (2015). Measuring Great Jobs and Great Lives: The Gallup-Purdue Index. *IEEE Computer*, *48*(1), 66-69.

Darby, N. M., & Rashid, A. M. (2017). Critical Thinking Disposition: The Effects of Infusion Approach in Engineering Drawing. *Journal of Education and Learning*, *6*(3), 305-311.

Davis, A. L. (2013). Using instructional design principles to develop effective information literacy instruction: The ADDIE model. *College & Research Libraries News*, *74*(4), 205-207.

De Lia, E., & Fredericks, E. (2005). *From cross purposes to cooperation: The ten factors that unify a cross-functional team*. iUniverse.

Department of Statistics Malaysia. (2016). *Labour Force and Social Statistics*. Retrieved from: https://www.statistics.gov.my/index.php?r=column/cone&menu_id=bVNhYUZZbmd3VDMrcGNxUlNKSWNwUT09 rate-trend-midf-research-9495

Department of Statistics, Malaysia (2011). Statistics Labour Force. Ministry of Human Resources

Dewey, J. (1933). How we think: A restatement of the relation of reflective thinking to the educative process (1910), revised edition. *Boston: Heath*.

Dewey, J. (1938). *Experience and education.* New York, NY: Macmillan

DiIorio, C. K. (2006). *Measurement in health behavior: methods for research and evaluation* (Vol. 1). John Wiley & Sons.

Djaouti, D., Alvarez, J., Jessel, J. P., & Rampnoux, O. (2011). Origins of serious games. In *Serious games and edutainment applications* (pp. 25-43). Springer, London.

Donnelly, D. F., Linn, M. C., & Ludvigsen, S. (2014). Impacts and characteristics of computer-based science inquiry learning environments for precollege students. *Review of Educational Research*, *84*(4), 572-608.

dos Santos, A. D., & Fraternali, P. (2015, December). A comparison of methodological frameworks for digital learning game design. In *International conference on games and learning alliance* (pp. 111-120). Springer, Cham.

Dotson, L., & Foley, V. (2017). Common Core, Socioeconomic Status, and Middle Level Student Achievement: Implications for Teacher Preparation Programs in Higher Education. *Journal of Education and Learning*, *6*(4), 294-302.

Downe, A. G., Loke, S. P., Ho, J. S. Y., & Taiwo, A. A. (2012). Corporate talent needs and availability in Malaysian service industry. *International Journal of Business and Management*, *7*(2), 224.

Dunlop, L., Clarke, L., & McKelvey-Martin, V. (2019). Free-choice learning in school science: A model for collaboration between formal and informal science educators. *International Journal of Science Education, Part B*, *9*(1), 13-28.

Ecker, M., Müller, W., & Zylka, J. (2011). Game-based learning design patterns: An approach to support the development of better educational games. In P. Felicia (Ed.), *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches* (pp. 137-152). doi: 10.4018/978-1-60960-495-0.ch007

Economic Planning Unit (2015) Eleventh Malaysia plan 2016–2020: anchoring growth on people. Prime Minister's Department, Putrajaya.

Ehren, M. and Hatch, T. (2013). Responses of schools to accountability systems using multiple measures: the case of New York City elementary schools. *Educational Assessment, Evaluation and Accountability*, 25(4), pp.341-373.

Eller, R. (2015). Implementing Design Into Instruction : Intersections Between the Waterfall Model and Addie Model. *AU-eJournal of Interdisciplinary Research, 1*(1), 63–69.

El-Nasr, M. S., & Smith, B. K. (2006). Learning through game modding. *Computers in Entertainment*, *4*(1), 45–64. https://doi.org/10.1145/1111293.1111301

Ennis, R. (2001). An Outline of Goals for A Critical Thinking Curriculum and Its Assessment in Costa, A.(Ed.) Developing Minds: A Resource Book for Teaching Thinking. *Alexandria, VA: Association for Supervision and Curriculum Development*.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational researcher*, *18*(3), 4-10.

Ennis, R. H., Millman, J., & Tomko, T. N. (1985). *Cornell critical thinking tests level X & level Z: Manual*. Pacific Grove, CA: Midwest Publications.

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & education*, *67*, 156-167.

Eye, A. M., Gilb, M. L., & Hicks, V. M. (2013). *A policy analysis addressing 21st century learning* (Doctoral dissertation, Saint Louis University).

Facione, N. C., & Facione, P. A. (2008). *Critical thinking and clinical reasoning in the health sciences: An international multidisciplinary teaching anthology* (pp. 1-13). California Academic Press.

Facione, P. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations*. ERIC Document Reproduction Service No. ED315423.

Fadhlullah, A., & Ahmad, N. (2017). Thinking outside of the box: Determining students' level of critical thinking skills in teaching and learning. *Asian Journal of University Education*, *13*(2), 51-70.

Fadli, A., & Irwanto. (2020). The effect of local wisdom-based ELSII learning model on the problem solving and communication skills of pre-service islamic teachers. *International Journal of Instruction*, *13*(1), 731–746.

Fadzil, H. M., & Saat, R. M. (2014). Enhancing STEM Education during School Transition: Bridging the Gap in Science Manipulative Skills. *Eurasia Journal of Mathematics, Science & Technology Education*, *10*(3).

Falloon, G. (2017). Mobile devices and apps as scaffolds to science learning in the primary classroom. *Journal of Science Education and Technology*, *26*(6), 613-628.

Farhan, E. (2015). Gamejam as Design Method.

Fazaniza E (2016) Fresh grads face tough job market. The sun daily, http://www.thesundaily.my/news/1660165

Fieeq. (2011, July 14). i'm sleeeeeeepy ! but i can't sleep. [Personal blog]. Retrieved from http://feefuad.blogspot.com/2011/07/im-sleeeeeeepy-but-i-cant-sleep.html

Filsecker, M., & Hickey, D. T. (2014). A multilevel analysis of the effects of external rewards on elementary students' motivation, engagement and learning in an educational game. *Computers & Education*, *75*, 136-148.

Fiorella, L., & Mayer, R. E. (2012). Paper-based aids for learning with a computer-based game. *Journal of Educational Psychology*, *104*(4), 1074.

Fisher, A. (2011). *Critical thinking: An introduction*. Cambridge University Press.

Flores, K. L., Matkin, G. S., Burbach, M. E., Quinn, C. E., & Harding, H. (2012). Deficient critical thinking skills among college graduates: Implications for leadership. *Educational Philosophy and Theory*, *44*(2), 212-230.

Foster, A., & Shah, M. (2015). The ICCE framework: Framing learning experiences afforded by games. *Journal of Educational Computing Research*, *51*(4), 369-395.

Franco, A., Marques Vieira, R., & Tenreiro-Vieira, C. (2018). Educating for critical thinking in university: The criticality of critical thinking in education and everyday life. *Essachess. Journal for Communication Studies*, *11*(2), 22.

Friedman, T. L. (2005). *The world is flat: A brief history of the twenty-first century*. Macmillan.

Fuad, M., Deb, D., Etim, J., & Gloster, C. (2018). Mobile response system: a novel approach to interactive and hands-on activity in the classroom. *Educational technology research and development*, *66*(2), 493-514.

Fullerton, T., Swain, C., & Hoffman, S. (2004). Game Design Workshop–Designing, Prototyping, and Playtesting Games, 2004. *Google Scholar Google Scholar Digital Library Digital Library*.

Galia, J. (2008). Copyright© 2008 International Association for the Evaluation of Educational Achievement (IEA) Published December 2008, Revised August 2009 TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades Michael O. Martin, Ina VS Mullis, Pierre Foy in collaboration with John F. Olson, Ebru Erberber, Corinna Preuschoff.

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE)*, *1*(1), 20-20.

Gelder, T. (2001, December). How to improve critical thinking using educational technology. In *Meeting at the crossroads: Proceedings of the 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 539-548).

Gelerstein, D., Del Rio, R., Nussbaum, M., Chiuminatto, P., & López, X. (2016). Designing and implementing a test for measuring critical thinking in primary school. *Thinking Skills and Creativity*, *20*, 40-49.

George, D., & Mallery, P. (2019). IBM SPSS statistics 25 step by step: A simple guide and reference (15th ed.). New York: Routledge.

Gillispie, L., Martin, F., & Parker, A. M. (2010). Effects of a 3-D video game on middle school student achievement and attitude in mathematics. *Electronic Journal of Mathematics and Technology, 4*, 1933–2823.

Gillispie, L., Martin, F., & Parker, M. (2009, March). Effects of the dimension-M 3D video gaming experience on middle school student achievement and attitude in mathematics. In *Society for Information Technology & Teacher Education International Conference* (pp. 1462-1469). Association for the Advancement of Computing in Education (AACE).

Ginsburg, A., & Smith, M. S. (2016). Do randomized controlled trials meet the "gold standard". *American Enterprise Institute. Retrieved March*, *18*, 2016.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). Highlights from TIMSS 2007: Mathematics and Science Achievement of US Fourth-and Eighth-Grade Students in an International Context. NCES 2009-001. *National Center for Education Statistics*.

Graf-Webster, E. (2011). Standardized Testing: Good or Bad for Assessment of Teacher Performance, Assessment of the Education System?.

Greiff, S., Niepel, C., & Wüstenberg, S. (2015). 21st century skills: International advancements and recent developments. *Thinking Skills and Creativity*, *100*(18), 1-3.

Guilford, J. P., & Fruchter, B. (1973). Fundamental statistics in education and psychology.

Gurria, A. (2018). PISA 2018 Results COBINED EXECUTIVE SUMMARIES VOLUME I, II, & III (p. 18).

Hackathorn, J., Solomon, E. D., Blankmeyer, K. L., Tennial, R. E., & Garczynski, A. M. (2011). Learning by Doing: An Empirical Study of Active Teaching Techniques. *Journal of Effective Teaching*, *11*(2), 40-54.

Hainey, T., Connolly, T. M., Boyle, E. A., Wilson, A., & Razak, A. (2016). A systematic literature review of games-based learning empirical evidence in primary education. *Computers & Education*, *102*, 202-223.

Hainey, T., Connolly, T. M., Stansfield, M., & Boyle, E. A. (2011). Evaluation of a game to teach requirements collection and analysis in software engineering at tertiary education level. *Computers & Education*, *56*(1), 21-35.

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.

Hamari, J., & Koivisto, J. (2014). Measuring flow in gamification: Dispositional flow scale-2. *Computers in Human Behavior*, *40*, 133-143.

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in human behavior*, *54*, 170-179.

Hanapi, Z., & Nordin, M. S. (2014). Unemployment among Malaysia graduates: Graduates' attributes, lecturers' competency and quality of education. *Procedia-Social and Behavioral Sciences*, *112*(2014), 1056-1063.

Harris, C. (2010). Going deep. *School Library Journal, 56*(11), 38-40.

Harteveld, C., Guimarães, R., Mayer, I. S., & Bidarra, R. (2010). Balancing play, meaning and reality: The design philosophy of LEVEE PATROLLER. *Simulation & Gaming*, *41*(3), 316-340.

Hasegawa, T., Koshino, M., & Ban, H. (2015). An English vocabulary learning support system for the learner's sustainable motivation. *SpringerPlus*, *4*(1), 99.

Heiniger, L. E., Clark, G. I., & Egan, S. J. (2018). Perceptions of Socratic and non-Socratic presentation of information in cognitive behaviour therapy. *Journal of Behavior Therapy and Experimental Psychiatry*, *58*, 106–113.

Hepner, M. R. (2012). *Evaluating the critical thinking skills and academic characteristics of undergraduate students at two post-secondary institutions utilizing two different curriculum models* (Doctoral dissertation, Lindenwood University).

Heyman, G. D., & Legare, C. H. (2005). Children's evaluation of sources of information about traits. *Developmental Psychology*, *41*(4), 636.

Honey, M. A., & Hilton, M. L. (2011). Learning science through computer games. *National Academies Press, Washington, DC*.

Hong, J. C., Hwang, M. Y., Lu, C. H., Cheng, C. L., Lee, Y. C., & Lin, C. L. (2009). Playfulness-based design in educational games: A perspective on an evolutionary contest game. *Interactive Learning Environments*, *17*(1), 15-35.

Hooshyar, D., Pedaste, M., Saks, K., Leijen, Ä., Bardone, E., & Wang, M. (2020). Open learner models in supporting self-regulated learning in higher education: A systematic literature review. *Computers & Education*, *154*, 103878.

Hopson, M. H., Simms, R. L., & Knezek, G. A. (2001). Using a technology-enriched environment to improve higher-order thinking skills. *Journal of Research on Technology in education*, *34*(2), 109-119.

Horn, R. E. (1995). The story of the guide to simulations/games for education and training. *Simulation & Gaming*, *26*(4), 471-480.

Hossain, M. I., Yagamaran, K. S., Afrin, T., Limon, N., Nasiruzzaman, M., & Karim, A. M. (2018). Factors Influencing Unemployment among Fresh Graduates: A Case Study in Klang Valley, Malaysia. *International Journal of Academic Research in Business & Social Sciences*, *8*, 1494-1507.

Hou, H. & Li, M. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Computers In Human Behavior*, *30*, 29–38.

Hovious, A. S. (2015). Digital games for 21st-century learning. Teacher Librarian, 42(5), 34–38

Hsu, T. C., Lee-Hsieh, J., Turton, M. A., & Cheng, S. F. (2014). Using the ADDIE model to develop online continuing education courses on caring for nurses in Taiwan. *The Journal of Continuing Education in Nursing*, *45*(3), 124-131.

Hung, C. Y., Sun, J. C. Y., & Yu, P. T. (2015). The benefits of a challenge: student motivation and flow experience in tablet-PC-game-based learning. *Interactive Learning Environments*, *23*(2), 172-190.

Hung, H. T., Yang, J. C., Hwang, G. J., Chu, H. C., & Wang, C. C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education*, *126*, 89-104.

Hwa, S. P. (2018). Pedagogical change in mathematics learning: Harnessing the power of digital game-based learning. *Journal of Educational Technology & Society*, *21*(4), 259-276.

Hwa, Y. Y. (2016). From drills to skills? Cultivating critical thinking, creativity, communication, and collaboration through Malaysian schools. *Working paper*.

Hwang, G. J., & Chen, C. H. (2017). Influences of an inquiry-based ubiquitous gaming design on students' learning achievements, motivation, behavioral patterns, and tendency towards critical thinking and problem solving. *British Journal of Educational Technology*, *48*(4), 950-971.

Hwang, G. J., & Wu, P. H. (2012). Advancements and trends in digital game-based learning research: a review of publications in selected journals from 2001 to 2010. *British Journal of Educational Technology*, *43*(1), E6-E10.

Hwang, G. J., Sung, H. Y., Hung, C. M., Huang, I., & Tsai, C. C. (2012). Development of a personalized educational computer game based on students' learning styles. *Educational Technology Research and Development*, *60*(4), 623-638.

Hwang, G. J., Sung, H. Y., Hung, C. M., Yang, L. H., & Huang, I. (2013a). A knowledge engineering approach to developing educational computer games for improving students' differentiating knowledge. *British journal of educational technology*, *44*(2), 183-196.

Hwang, G. J., Wu, P. H., & Chen, C. C. (2012). An online game approach for improving students’ learning performance in web-based problem-solving activities. *Computers & Education*, *59*(4), 1246-1256.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, *69*, 121-130.

Hwang, G., Chiu, L., & Chen, C. (2015). A contextual game-based learning approach to improving students' inquiry-based learning performance in social studies courses. *Computers & Education, 81*, 13–25.

Hwang, G.-J., & Chang, H.-F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education, 56*, 1023–1031.

Iqbal, H. M., Bibi, F., & Iqbal, N. (2015). Exploring factors affecting grade VIII students’ attainment in science: options for reforms. *American Journal of Educational Science*, *1*(4), 210-217.

Ismail, N. S., Harun, J., Zakaria, M. A. Z. M., & Salleh, S. M. (2018). The effect of mobile problem-based learning application DicScience PBL on students’ critical thinking. *Thinking Skills and Creativity*, *28*, 177-195.

Ismail, R., Yussof, I., & Sieng, L. W. (2011). Employers’ perceptions on graduates in Malaysian services sector. *International Business Management*, *5*(3), 184-193.

Ismail, W., Saifuddin, W. O. A., & Noraini, H. (2017, November). Critical thinking and problem solving skills among engineering students at public universities in Malaysia. In *2017 7th World Engineering Education Forum (WEEF)* (pp. 262-267). IEEE.

Jabbar, A. I. A., & Felicia, P. (2015). Gameplay engagement and learning in game-based learning: a systematic review. *Review of Educational Research, 85*(4), 740–779.

Jackson, G. T., & McNamara, D. S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology*, *105*(4), 1036.

Jonassen, D. H. (2012). Designing for decision making. *Educational technology research and development*, *60*(2), 341-359.

Kalonji, G. (2005). Capturing the imagination: High-priority reforms for engineering educators. *Educating the engineer of 2020: Adapting engineering education to the new century*, 146-450.

Kanbay, Y., & Okanlı, A. (2017). The effect of critical thinking education on nursing students’ problem-solving skills. *Contemporary nurse*, *53*(3), 313-321.

Kao, G. Y. M., Chiang, C. H., & Sun, C. T. (2017). Customizing scaffolds for game-based learning in physics: Impacts on knowledge acquisition and game design creativity. *Computers & Education*, *113*, 294-312.

Ke, F. (2016). Designing and integrating purposeful learning in game play: A systematic review. *Educational Technology Research and Development, 64(2)*, 219–244.

Kementerian Pendidikan Malaysia. (2013). *Malaysia Education Blueprint 2013-2025 (Preschool to Post-Secondary Education)*. Putrajaya, Malaysia: Kementerian Pendidikan Malaysia.

Kennedy, D. (2014). What can 'gamification' do for lawyers?. *ABA Journal*, *100*(3), 1.

Khalaf, B. K. (2018). Traditional and Inquiry-Based Learning Pedagogy: A Systematic Critical Review. *International Journal of Instruction*, *11*(4), 545-564.

Kim, H. Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative dentistry & endodontics*, *38*(1), 52-54.

Kim, M., Anthony, R., & Blades, D. (2014). Decision making through dialogue: A case study of analyzing preservice teachers' argumentation on socioscientific issues. *Research in Science Education*, *44*(6), 903-926.

King, D. L., Delfabbro, P. H., & Griffiths, M. D. (2011). The role of structural characteristics in problematic video game play: an empirical study. *International Journal of Mental Health and Addiction, 9*(3), 320–333.

Klimek, R., & Szwed, P. (2010). Formal analysis of use case diagrams. *Computer Science*, *11*, 115-131.

Knapp, N. F. (2019). The shape activity: Social constructivism in the psychology classroom. *Teaching of Psychology*, *46*(1), 87-91.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child development*, *76*(6), 1261-1277.

Kong, S. C. (2014). Developing information literacy and critical thinking skills through domain knowledge learning in digital classrooms: An experience of practicing flipped classroom strategy. *Computers & Education*, *78*, 160-173.

Kong, S. C. (2015). An experience of a three-year study on the development of critical thinking skills in flipped secondary classrooms with pedagogical and technological support. *Computers & Education*, *89*, 16-31.

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kowalczyk, N., Hackworth, R., & Case-Smith, J. (2012). Perceptions of the use of critical thinking teaching methods. *Radiologic technology*, *83*(3), 226-236.

Ku, O., Chen, S. Y., Wu, D. H., Lao, A. C., & Chan, T. W. (2014). The effects of game-based learning on mathematical confidence and performance: High ability vs. low ability. *Journal of Educational Technology & Society*, *17*(3), 65-78.

Kuhn, D. (1991) *The Skills of Argument*. Cambridge: Cambridge University Press.

Lai, E. R. (2011). Critical thinking: A literature review. *Pearson's Research Reports*, *6*, 40-41.

Lai, H. M., Hsiao, Y. L., & Hsieh, P. J. (2018). The role of motivation, ability, and opportunity in university teachers' continuance use intention for flipped teaching. *Computers and Education*, *124*(November 2017), 37–50.

Lambert (2012). Twilight of the lecture. *Harvard Magazine*. Retrieved December 30, 2016, from http://harvardmagazine.com/2012/03/ twilight-of-the-lecture

Landsman, J., & Gorski, P. (2007). Countering Standardization. *Educational leadership*, *64*(8), 40-44.

Lang, J. (2014). How students learn from games. *The Chronicle of higher education*, *61*(3), B34-B35.

Lee, Y. F., Lin, C. J., Hwang, G. J., Fu, Q. K., & Tseng, W. H. (2021). Effects of a mobile-based progressive peer-feedback scaffolding strategy on students' creative thinking performance, metacognitive awareness, and learning attitude. *Interactive Learning Environments*, 1-17.

Leen, C. C., Hong, H., Kwan, F. F. H., & Ying, T. W. (2014). Creative and critical thinking in Singapore schools. *Singapore: Nanyang Technological University*.

Leys, C., Klein, O., Dominicy, Y., & Ley, C. (2018). Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance. *Journal of Experimental Social Psychology*, *74*(September 2017), 150–156. https://doi.org/10.1016/j.jesp.2017.09.011

Li, F. Y., Hwang, G. J., Chen, P. Y., & Lin, Y. J. (2021). Effects of a concept mapping-based two-tier test strategy on students' digital game-based learning performances and behavioral patterns. *Computers and Education*, *173*(July), 104293.

Li, M. C., & Tsai, C. C. (2013). Game-based learning in science education: A review of relevant research. *Journal of Science Education and Technology*, *22*(6), 877-898.

Liao, C. W., Chen, C. H., & Shih, S. J. (2019). The interactivity of video and collaboration for learning achievement, intrinsic motivation, cognitive load, and behavior patterns in a digital game-based learning environment. *Computers & Education*, *133*, 43-55.

Licorish, S. A., Owen, H. E., Daniel, B., & George, J. L. (2018). Students' perception of Kahoot!'s influence on teaching and learning. *Research and Practice in Technology Enhanced Learning*, *13*(1), 1-23.

Lim, S. L. (2011). *Social networks and collaborative filtering for large-scale requirements elicitation* (Doctoral dissertation, University of New South Wales).

Lin, H. W., & Lin, Y. L. (2014). Digital educational game value hierarchy from a learners' perspective. *Computers in Human Behavior*, *30*, 1-12.

Linehan, C., Kirman, B., Lawson, S., & Chan, G. (2011, May). Practical, appropriate, empirically-validated guidelines for designing educational games. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1979-1988). ACM.

Liu, C., & Long, F. (2014, January). The discussion of traditional teaching and multimedia teaching approach in college English teaching. In *2014 International Conference on Management, Education and Social Science (ICMESS 2014)* (pp. 31-33). Atlantis Press.

Liu, C., Cheng, Y., & Huang, C. (2011). The effect of simulation games on the learning of computational problem solving. *Computers & Education*, *57*(3), 1907–1918.

Liu, H., Wang, L., & Koehler, M. J. (2019). Exploring the intention-behavior gap in the technology acceptance model: A mixed-methods study in the context of foreign-language teaching in China. *British Journal of Educational Technology*, *50*(5), 2536-2556.

López-Cabrales, A., & Bornay-Barrachina, M. (2019). Decision Making: International Managerial Skills. In *Managerial Competencies for Multinational Businesses* (pp. 64-81). IGI Global.

Lowther, D. L., Inan, F. A., Strahl, J. D., & Ross, S. M. (2012). Do one-to-one initiatives bridge the way to 21st century knowledge and skills?. *Journal of Educational Computing Research*, *46*(1), 1-30.

Luca Chittaro. (2016). Designing Serious Games for Safety Education: "Learn to Brace" versus Traditional Pictorials for Aircraft Passengers. *Ieee Transactions on Visualization and Computer Graphics*, *22*, 1527–1539.

Magee, R. G., & Jones, B. D. (2012). An instrument to assess beliefs about standardized testing: Measuring the influence of epistemology on the endorsement of standardized testing. *Australian Journal of Educational & Developmental Psychology*, *12*, 71-82.

Malaysia, K. P. (2006). Pelan Induk Pembangunan Pendidikan 2006-2010. *Putrajaya: Kementerian Pelajaran Malaysia*.

Malaysian Education Act, (1996). Ministry of Education Malaysia, Kuala Lumpur.

Martin, M. O., Mullis, I. V. S., & Foy, P. (2007). TIMSS 2007 International Science Report. In *Science*.

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). TIMSS 2015 international results in science. Retrieved from Boston College, TIMSS & PIRLS International

Study Center website: http://timssandpirls.bc.edu/timss2015/international-results/

Martin, M. O., Mullis, I. V. S., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 international results in science.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Martin, M. O., Mullis, I. V., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 International Results in Science*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.

Marx, G. (2011). *Sixteen Trends... Their Profound Impact on Our Future*. Educational Research Service. 1001 North Fairfax Street Suite 500, Alexandria, VA 22314.

Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research*, *42*(3), 241-265.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. Penguin.

McGuinness, C. (2005). Teaching thinking: Theory and practice. In *BJEP monograph series II, number 3-pedagogy-teaching for learning* (Vol. 107, No. 126, pp. 107-126). British Psychological Society.

McGuire, L. A. (2010). *Improving Student Critical Thinking and Perceptions of Critical Thinking through Direct Instruction in Rhetorical Analysis*. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.

McMahon, G. (2009). Critical thinking and ICT integration in a Western Australian secondary school. *Journal of Educational Technology & Society*, *12*(4), 269-281.

McPECK, J. (1990). Teaching Critical Thinking New York: Routledge.

McPhail, G. (2016). The fault lines of recontextualisation: The limits of constructivism in education. *British Educational Research Journal*, *42*(2), 294-313.

Meluso, A., Zheng, M., Spires, H., & Lester, J. (2012). Enhancing 5th graders' science content knowledge and self-efficacy through game-based learning. *Computers & Education*, *59*(2), 497–504.

Meng, C. C., Idris, N., & Eu, L. K. (2014). Secondary Students' Perceptions of Assessments in Science, Technology, Engineering, and Mathematics (STEM). *Eurasia Journal of Mathematics, Science & Technology Education*, *10*(3).

Menn, D. (1993, October). Multimedia in education. *PC World*, M52-M60

Michael, D. R., & Chen, S. L. (2005). *Serious games: Games that educate, train, and inform*. Boston, Mass: Cengage Learning PTR.

Michel, N., Cater III, J. J., & Varela, O. (2009). Active versus passive teaching styles: An empirical study of student learning outcomes. *Human resource development quarterly*, *20*(4), 397-418.

Mikropolous, T. A., & Natsis, A. (2011). Educational virtual environments: a ten-year review of empirical research. *Computers & Education, 56*, 769–780. doi:10.1016/j.compedu.2010.10.020.

Miri, B., David, B. C., & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in science education*, *37*(4), 353–369.

Mohamad Johdi Salleh, A., & Hatta, M. (2017). *Anjuran: Pejabat Pendidikan Daerah Hilir Perak and Bagan Datuk*.

Mohtar, L. E., Halim, L., Samsudin, M. A., & Ismail, M. E. (2019). Non-cognitive factors influencing science achievement in Malaysia and Japan: An analysis of TIMSS 2015. *EURASIA Journal of Mathematics, Science and Technology Education*, *15*(4), em1697.

Mojumder, F., & Keast, S. (2018). 'School Science Is Irrelevant': Why Still Do Science? A Case Study on Secondary Students in Bangladesh. In *Engaging in Educational Research* (pp. 201-220). Springer, Singapore.

Molenda, M. (2015). In search of the elusive ADDIE model. *Performance Improvement, 54*(2), 40-42.

Molenda, M., Reigeluth, C. M., & Nelson, L. M. (2003). Instructional design. In (L. Nadel, Ed.) Encyclopedia of cognitive science.

Moon, M. K., Jang, S. K., & Kim, T. Y. (2011). A computer-assisted learning model based on the digital game exponential reward system. *The Turkish Online Journal of Educational Technology, 10*(1), 1–14.

Morrison, G. R., Ross, S. J., Morrison, J. R., & Kalman, H. K. (2019). *Designing effective instruction*. John Wiley & Sons.

Mowbray, F. I., Fox-Wasylyshyn, S. M., & El-Masri, M. M. (2019). Univariate Outliers: A Conceptual Overview for the Nurse Researcher. *The Canadian Journal of Nursing Research = Revue Canadienne de Recherche En Sciences Infirmieres*, *51*(1), 31–37. https://doi.org/10.1177/0844562118786647

Mullis, I. V. (2008). *TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. TIMSS & PIRLS International Study Center.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). TIMSS 2015 international results in mathematics. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2015/international-results/

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science.* Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/timss2019/international-results/

Mullis, I. V., & Martin, M. O. (2008). Overview of TIMSS 2007. *Chestnut Hill, MA: TIMSS & PIRLS*.

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.

Munassar, N. M. A., & Govardhan, A. (2011). Comparison between traditional approach and object-oriented approach in software engineering development. *International Journal of Advanced Computer Science and Applications*, *2*(6).

Murugan, A., & Rajoo, L. (2013). Students' perceptions of mathematics classroom environment and mathematics achievement: A study in Sipitang, Sabah, Malaysia. In *International Conference on Social Science Research, Penang, Malaysia*.

Muruganantham, G. (2015). Developing of E-content package by using ADDIE model. *International Journal of Applied Research*, *1*(3), 52-54.

Mustapha, R. B., & Greenan, J. P. (2002). The role of vocational education in economic development in Malaysia: educators' and employers' perspectives.

Navarrete, C. C. (2014, June). Digital literacies and 21st century skills: the students' game design and development experience. In *EdMedia+ Innovate Learning* (pp. 1177-1197). Association for the Advancement of Computing in Education (AACE).

Nazron, M. A., Lim, B., & Nga, J. L. (2017). Soft skills attributes and graduate employability: A case in Universiti Malaysia Sabah. *Malaysian Journal of Business and Economics (MJBE)*.

Ng, K. T., Lay, Y. F., Areepattamannil, S., Treagust, D. F., & Chandrasegaran, A. L. (2012). Relationship between affect and achievement in science and mathematics in Malaysia and Singapore. *Research in Science & Technological Education*, *30*(3), 225-237.

Ngoo, Y. T., Tiong, K. M., & Pok, W. F. (2015). Bridging the gap of perceived skills between employers and accounting graduates in Malaysia. *American Journal of Economics*, *5*(2), 98-104.

Nicholson, S. (2014). *Everyone plays at the library: Creating great gaming experiences for all ages*. Information Today, Incorporated.

Nikou, S., & Aavakare, M. (2021). An assessment of the interplay between literacy and digital Technology in Higher Education. *Education and Information Technologies*, 1-23.

Nooriah, Y., & Zakiah, J. (2017). Development of graduates employability: The role of university and challenges. *Jurnal Personalia Pelajar*, *20*(1), 15-32.

Nygren, E., Sutinen, E., Blignaut, A. S., Laine, T. H., & Els, C. J. (2012). Motivations for play in the UFractions mobile game in three countries. *International Journal of Mobile and Blended Learning (IJMBL)*, *4*(2), 30-48.

OECD (2016), "PISA 2015 Results in Focus", *PISA in Focus*, No. 67, OECD Publishing, Paris, https://doi.org/10.1787/aa9237e6-en.

OECD. (2017). *PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving*. OECD Publishing.

OECD. (2018). About- Pisa. Retrieved from https://www.oecd.org/pisa/aboutpisa/

OECD. (2019). *Malaysia What 15-year-old students in Malaysia know and can do*. http://www.oecd.org/pisa/publications/PISA2018_CN_MYS.pdf

OECD. (2019a). *Singapore What 15-year-old students in Malaysia know and can do.* https://www.oecd.org/pisa/publications/PISA2018_CN_SGP.pdf

Oliveira, M., Ribeiro, L., Cota, É., Duarte, L. M., Nunes, I., & Reis, F. (2015, September). Use case analysis based on formal methods: an empirical study. In *International Workshop on Algebraic Development Techniques* (pp. 110-130). Springer, Cham.

Oliver, R. (2008). Engaging first year students using a web-supported inquiry-based learning setting. *Higher Education, 55(3)*, 285–301.

Osborne, J., & Collins, S. (2001). Pupils' views of the role and value of the science curriculum: a focus-group study. *International journal of science education*, *23*(5), 441-467.

Ott, M., Popescu, M. M., Stanescu, I. A., & de Freitas, S. (2013). Game-enhanced learning: Preliminary thoughts on curriculum integration. In *New pedagogical approaches in game enhanced learning: Curriculum integration* (pp. 1-19). IGI Global.

Owen, S., Dickson, D., Stanisstreet, M., & Boyes, E. (2008). Teaching physics: Students' attitudes towards different learning activities. *Research in Science & Technological Education*, *26*(2), 113-128.

Ozkan-Akan, S. (2003). *Teachers' perceptions of constraints on improving student thinking in high schools.* Masters Thesis, Middle East Technical University, Departmentl of Social Sciences. Retrieved November 13, 2013.

Pacheco, C., García, I., & Reyes, M. (2018). Requirements elicitation techniques: a systematic literature review based on the maturity of the techniques. *IET Software*, *12*(4), 365-378.

Palmer, A. W. (2016). *Higher-order thinking skills in digital games* (Doctoral dissertation, Azusa Pacific University).

Papanastasiou, G. P., Drigas, A. S., & Skianis, C. (2017). Serious Games in Preschool and Primary Education: Benefits And Impacts on Curriculum Course Syllabus. *International Journal of Emerging Technologies in Learning*, *12*(1).

Pappa, D., & Pannese, L. (2010, September). Effective design and evaluation of serious games: The case of the e-VITA project. In *World Summit on Knowledge Society* (pp. 225-237). Springer, Berlin, Heidelberg.

Parkman, S., Litz, D., & Gromik, N. (2018). Examining pre-service teachers' acceptance of technology-rich learning environments: A UAE case study. *Education and Information Technologies*, *23*(3), 1253-1275.

Partnership for 21st Century Skills. (2009). *Learning environments: A 21st century skills implementation guide.*

Partnership for 21st Century Skills. (2011). Framework for 21st Century learning Accessed 16.02.20 http://www.p21.org/storage/documents/1.__p21_framework_2-pager.pdf

Partnership for 21st Century Skills. (2019). *Framework for 21st Century Learning A unified vision for learning to ensure student success in a world where change is constant and learning never stops.*

Paul, R. (2005). The state of critical thinking today. *New directions for community colleges*, *2005*(130), 27-38.

Paul, R., & Elder, L. (2007). Critical thinking: The art of Socratic questioning. *Journal of developmental education*, *31*(1), 36.

Paul, R., Elder, L., & Bartell, T. (1997). California teacher preparation for instruction in critical thinking: Research findings and policy recommendations. State of California, California Commission on Teacher Credentialing, Sacramento

Paul, R., Elder, L., & Bartell, T. (1997a). A brief history of the idea of critical thinking. *Retrieved July*, *10*, 2006.

Perinetti, G. (2018). StaTips Part IV: Selection, interpretation and reporting of the intraclass correlation coefficient. *South European Journal of Orthodontics and Dentofacial Research*, *5*(1), 3–5. https://doi.org/10.5937/sejodr5-17434

Perini, S., Luglietti, R., Margoudi, M., Oliveira, M., & Taisch, M. (2018). Learning and motivational effects of digital game-based learning (DGBL) for manufacturing education–The Life Cycle Assessment (LCA) game. *Computers in Industry*, *102*, 40-49.

Perrotta, C., Featherstone, G., Aston, H., & Houghton, E. (2013). Game-based learning: Latest evidence and future directions. *Slough: NFER*.

Perry, D. K. (2014). Exploring critical thinking skills among undergraduate agriculture education and studies students.

Persky, A. M., Medina, M. S., & Castleberry, A. N. (2019). Developing critical thinking skills in pharmacy students. *American Journal of Pharmaceutical Education*, *83*(2).

Phang, F. A., Khamis, N., Nawi, N. D., & Pusppanathan, J. (2020). TIMSS 2019 Science Grade 8: Where is Malaysia standing?. *Journal of Engineering Education*, *4*(2), 37-43.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures. (Trans A. Rosin)*. Viking.

Pillow, B. H. (2008). Development of children's understanding of cognitive activities. *The Journal of Genetic Psychology*, *169*(4), 297-321.

Power, J., Lynch, R., & McGarr, O. (2020). Difficulty and self‐efficacy: An exploratory study. *British Journal of Educational Technology*, *51*(1), 281-296.

Prensky, M. (2001). Fun, play and games: What makes games engaging. *Digital game-based learning*, *5*(1), 5-31.

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment (CIE)*, *1*(1), 21-21.

Prensky, M. (2007). “Don’t bother me Mom, I’m learning!”: how computer and video games are preparing your kids for twenty-first century success and how you can help! *Minnesota Paragon House*, 27–51.

Procci, K., Singer, A. R., Levy, K. R., & Bowers, C. (2012). Measuring the flow experience of gamers: An evaluation of the DFS-2. *Computers in Human Behavior*, *28*(6), 2306-2312.

Provasnik, S. D., Malley, L., Stephens, M., Landeros, K., Perkins, R., & Tang, J. H. (2016). Highlights from TIMSS and TIMSS Advance 2015. *Mathematics Framework*, *1*, 1-47.

Qian, M., & Clark, K. R. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in Human Behavior*, *63*, 50-58.

Rankin, Y. A. (2008). *Design and evaluation of massive multiplayer online role playing games that facilitate second language acquisition* (Doctoral dissertation, Northwestern University).

Rashid, R. A., & Hashim, R. A. (2008). The relationship between critical thinking and language proficiency of Malaysian undergraduates.

Razak, M. I. M., Ahmad, I., & De Mello, G. (2014). *Factors Influencing Unemployment among Graduates in Malaysia*. LAP LAMBERT Academic Publishing.

Renzulli, J. S. (2008). Engagement is the answer. *Education Week*, *27*(43), 30-31.

Reynolds, S. W. (2016). *Determining and exploring teachers' perceptions on the barriers to teaching critical thinking in the classroom: A survey study* (Doctoral dissertation).

Richey, R. C., Klein, J. D., & Tracey, M. W. (2011). *The instructional design knowledge base: Theory, research, and practice*. Routledge.

Rutenberg, J. (2003). A new attempt to monitor media content. *New York Times*, C2.

Sadera, J. R. N., Torres, R. Y. S., & Rogayan Jr, D. V. (2020). Challenges Encountered by Junior High School Students in Learning Science: Basis for Action Plan. *Universal Journal of Educational Research*, *8*(12A), 7405-7414.

Said, M., Mohsin, M., & Yunus, N. (2008). Halangan-halangan kepada usaha memupuk kreativiti di kalangan pelajar.

Salleh, A. H. (2006). Shifts in policies and approaches vis-a-vis English Language education in Malaysia. *English in Southeast Asia: Prospects, perspectives and possibilities*, 1-8.

Salute, C. (2015). *Digital Orientation and the Efficacy of Game-Based Training Platforms on Engagement and Knowledge Retention*. Hofstra University.

Samara, F. (2015). Factors influencing students' choice of elective science courses: A case study from the American University of Sharjah. *Open Journal of Social Sciences*, *3*(08), 93.

Sandberg, J., Maris, M., & Hoogendoorn, P. (2014). The added value of a gaming context and intelligent adaptation for a mobile learning application for vocabulary learning. *Computers & Education*, *76*, 119-130.

Sandholtz, J. H., Ogawa, R. T., & Scribner, S. P. (2004). Standards Gaps: Unintended Consequences of Local Standards-Based Reform. *Teachers college record*, *106*(6), 1177-1202.

Sardareh, S. A., & Saad, M. R. M. (2013). Malaysian Primary School ESL Teachers' Questions during Assessment for Learning. *English Language Teaching*, *6*(8), 1-9.

Sarneabat, M., & Kabir, M. (2018). Development of an android based mobile application to collect gas meter reading with appropriate customers approval.

Sauvé, L., Renaud, L., Kaufman, D., & Marquis, J. S. (2007). Distinguishing between games and simulations: A systematic review. *Journal of Educational Technology & Society*, *10*(3), 247-256.

Sawyer, B. (2007, March). The" serious games" landscape. In *Instructional & Research Technology Symposium for Arts, Humanities and Social Sciences, Camden, USA*

Sawyer, B., & Rejeski, D. (2002). Serious games: Improving public policy through game-based learning and simulation.

Schleicher, A. (2018). Insights and interpretations. *Pisa 2018*, *10*.

Schober, P., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia and Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Schrader, C. (2019). Creating avatars for technology usage: Context matters. *Computers in Human Behavior*, *93*, 219-225.

Scriven, M., & Paul, R. (2007). Defining Critical Thinking. The critical thinking community: foundation for critical thinking. *Retrieved July*, *2*, 2017.

Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. John Wiley & Sons.

Seman, S. C., Yusoff, W. M. W., & Embong, R. (2017). Teachers challenges in teaching and learning for higher order thinking skills (HOTS) in primary school. *International Journal of Asian Social Science*, *7*(7), 534-545.

Sendurur, E. (2018). Students as information consumers: A focus on online decision making process. *Education and Information Technologies*, *23*(6), 3007-3027.

Seng, L. C. (2018). Malaysia Public Universities' Graduate Employability Policies: An Analysis of First Degree Graduates Unemployment and Underemployment Issues.

Shaffer, D. (2006). *How computer games help children learn (1st ed.)*. New York: Palgrave Macmillan.

Shaharom, A. (2004). Reading comprehension test as a measurement of critical thinking ability. *Unpublished Ph.D. dissertation, University Pendidikan Sultan Idris, Malaysia*.

Shan, P. L. M., Yunus, M. M., & Mohamad, M. (2018). The Malaysian education blueprint 2013 and its effects on English language teaching in Malaysia. *Asian EFL Journal*, *20*(7), 66–80.

Sheldon, K. M., & Biddle, B. J. (1998). Standards, accountability, and school reform: Perils and pitfalls. *Teachers college record*, *100*, 164-180.

Shelton, S. A., & Brooks, T. (2019). "We Need to Get These Scores Up": A Narrative Examination of the Challenges of Teaching Literature in the Age of Standardized Testing.

Shepherd, A., & Cosgrif, B. (1998). Problem-based learning: A bridge between planning education and planning practice. *Journal of planning education and research*, *17*(4), 348-357.

Shi, Y. R., & Shih, J. L. (2015). Game factors and game-based learning design model. *International Journal of Computer Games Technology, 2015*, 1–11.

Shibley, I., Amaral, K. E., Shank, J. D., & Shibley, L. R. (2011). Designing a blended course: Using ADDIE to guide instructional design. *Journal of College Science Teaching*, *40*(6).

Shih, J. L., Shih, B. J., Shih, C. C., Su, H. Y., & Chuang, C. W. (2010). The influence of collaboration styles to children's cognitive performance in digital problem-solving game "William Adventure": a comparative case study. *Computers & Education*, 1–12.

Siddiquah, A., & Salim, Z. (2017). The ICT facilities, skills, usage, and the problems faced by the students of higher education. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(8), 4987-4994.

Simões, J., Redondo, R. D., & Vilas, A. F. (2013). A social gamification framework for a K-6 learning platform. *Computers in Human Behavior*, *29*(2), 345-353.

Siti Rahayah Ariffin dan Nor Azaheen Abdul Hamid (2009). Profil Kemahiran Pemikiran Kritikal Antara Pelajar Aliran Sains dan Bukan Sains. 2009. Malaysian Education Deans' Council Journal. Vol 3, 121-135

Smith, M. L., & Fey, P. (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education*, *51*(5), 334-344.

Snyder, L. G., & Snyder, M. J. (2008). Teaching critical thinking and problem solving skills. *The Journal of Research in Business Education*, *50*(2), 90.

Squire, K. D. (2008). Video game–based learning: An emerging paradigm for instruction. *Performance Improvement Quarterly*, *21*(2), 7-36.

Squire, K., & Steinkuehler, C. (2005). Meet the gamers. *Library journal*, *130*(7), 38-41.

Sternberg, R. J. (2004). Four alternative futures for education in the United States: It's our choice. *School Psychology Review*, *33*(1), 67-77.

Stewart, C., Wall, A., & Marciniec, S. (2016, July). Mixed signals: Do college graduates have the soft skills that employers want?. In *Competition Forum* (Vol. 14, No. 2, p. 276). American Society for Competitiveness.

Sumirattana, S., Makanong, A., & Thipkong, S. (2017). Using realistic mathematics education and the DAPIC problem-solving process to enhance secondary school students' mathematical literacy. *Kasetsart Journal of Social Sciences*, *38*(3), 307-315.

Sun, C.-T., Wang, D.-Y., & Chan, H.-L. (2011). How digital scaffolds in games direct problem-solving behaviors. *Computers & Education, 57*, 2118–2125.

Sung, H. Y., & Hwang, G. J. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers & Education*, *63*, 43–51.

Sung, H. Y., & Hwang, G. J. (2018). Facilitating effective digital game-based learning behaviors and learning performances of students based on a collaborative knowledge construction strategy. *Interactive Learning Environments*, *26*(1), 118-134.

Susi, T., Johannesson, M., & Backlund, P. (2007). *Serious games – an overview* Technical Report HS- IKI –TR-07-001 28 p.

Tahir, R., & Wang, A. I. (2020). Codifying game-based learning: Development and application of leaguÊ framework for learning games. Academic Conferences and Publishing International.

Tan, K. J., Ismail, Z., & Abidin, M. (2018). A comparative analysis on cognitive domain for the Malaysian primary four textbook series. *EURASIA Journal of Mathematics, Science and Technology Education*, *14*(4), 1273-1286.

Tan, L. J., Goh, H.-L. D., Ang, P. R., & Huan, S. V. (2013). Participatory evaluation of an educational game for social skills acquisition. *Computers & Education, 64*, 70–80. doi:10.1016/j.compedu.2013.01.006.

Taub, M., & Azevedo, R. (2018). Using Sequence Mining to Analyze Metacognitive Monitoring and Scientific Inquiry Based on Levels of Efficiency and Emotions during Game-Based Learning. *Journal of Educational Data Mining*, *10*(3), 1-26.

Taub, M., Sawyer, R., Lester, J., & Azevedo, R. (2020). The Impact of Contextualized Emotions on Self-Regulated Learning and Scientific Reasoning during Learning with a Game-Based Learning Environment. *International Journal of Artificial Intelligence in Education*, *30*(1), 97-120.

Thomson, S., De Bortoli, L., & Underwood, C. (2016). PISA 2015: A first look at Australia's results.

Ting, S. H., Marzuki, E., Chuah, K. M., Misieng, J., & Jerome, C. (2017). EMPLOYERS'VIEWS ON IMPORTANCE OF ENGLISH PROFICIENCY AND COMMUNICATION SKILL FOR EMPLOYABILITY IN MALAYSIA. *Indonesian Journal of Applied Linguistics*, *7*(2), 315-327.

Tiruneh, D. T., De Cock, M., Weldeslassie, A. G., Elen, J., & Janssen, R. (2017). Measuring Critical Thinking in Physics: Development and Validation of a Critical Thinking Test in Electricity and Magnetism. *International Journal of Science and Mathematics Education*, *15*(4), 663–682.

Tseng, F. C. (2011). Segmenting online gamers by motivation. *Expert Systems with Applications*, *38*(6), 7693-7697.

Tseng, F. C., & Teng, C. I. (2015). Online Gamers' Preferences for Online Game Charging Mechanisms: The Effect of Exploration Motivation. *International Journal of E-Business Research (IJEBR)*, *11*(1), 23-34.

U.S. Department of Education. (2015). *Fact sheet: Testing action plan* [Press release]. Retrieved https://www.ed.gov/news/press-releases/fact-sheet-testing-action-plan

Ugwuozor, F. O., Ede, M. O., Ifelunni, O. C., & Abiogu, G. C. (2020). Teachers' demographic variables as predictors of critical thinking skills of school children: implications for school counselling. *Glob J Health Sci*, *12*, 91-101.

Vaillancourt, O., & Egli, R. (2011). Placeholders beyond static art replacement. *Game engine gems*, *2*.

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE review*, *41*(2), 16.

Van Eck, R. (2015). Digital game-based learning: Still restless, after all these years. *EDUCAUSE review*, *50*(6), 13.

van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2020). Determinants of 21st-Century Skills and 21st-Century Digital Skills for Workers: A Systematic Literature Review. *SAGE Open*, *10*(1).

van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, *72*, 577–588. https://doi.org/10.1016/j.chb.2017.03.010

Vogel, B., Kurti, A., Milrad, M., Johansson, E., & Müller, M. (2014). Mobile inquiry learning in Sweden: Development insights on interoperability, extensibility and sustainability of the LETS GO software system. *Journal of Educational Technology & Society*, *17*(2), 43-57.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research*, *34*(3), 229-243.

Volante, L. (2004). Teaching to the Test: What Every Educator and Policy-Maker Should Know. *Canadian Journal of Educational Administration and Policy*.

Voskoglou, M. G., & Salem, A. B. M. (2020). Benefits and Limitations of the Artificial with Respect to the Traditional Learning of Mathematics. *Mathematics*, *8*(4), 611.

Wang, H. Y., Duh, H. B. L., Li, N., Lin, T. J., & Tsai, C. C. (2014). An investigation of university students' collaborative inquiry learning behaviors in an augmented reality simulation and a traditional simulation. *Journal of Science Education and Technology*, *23*(5), 682-691.

Watson, G., & Glaser, E. M. (1980). *Critical thinking appraisal: Manual*. Psychological Corporation, New York.

Wegrzecka-Kowalewski, E. (2018). *Critical Thinking in Intensive Language Programs for International Students in US Universities* (Doctoral dissertation, University of Pittsburgh).

Westera, W. (2019). Why and how serious games can become far more effective: Accommodating productive learning experiences, learner motivation and the monitoring of learning gains. *Journal of Educational Technology & Society*, *22*(1), 59-69.

Wilkinson, P. (2016). A brief history of serious games. In *Entertainment computing and serious games* (pp. 17-41). Springer, Cham.

Willingham, D. T. (2008). Critical thinking: Why is it so hard to teach?. *Arts Education Policy Review*, *109*(4), 21-32.

Wong, D. (2007). Beyond control and rationality: Dewey, aesthetics, motivation, and educative experiences. *Teachers College Record*, *109*(1), 192-220.

Wong, G. K. W., & Cheung, H. Y. (2018). Exploring children's perceptions of developing twenty-first century skills through computational thinking and programming. *Interactive Learning Environments*, 1-13.

Woo, J. C. (2014). Digital game-based learning supports student motivation, cognitive success, and performance outcomes. *Journal of Educational Technology & Society*, *17*(3), 291-307.

Wouters, P., & Van Oostendorp, H. (2017). Overview of instructional techniques to facilitate learning and motivation of serious games. In *Instructional techniques to facilitate learning and motivation of serious games* (pp. 1-16). Springer, Cham.

Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of educational psychology*, *105*(2), 249.

Wu, S. (2008). *Reducing unproductive learning activities in serious games for second language acquisition*. University of Southern California.

Wu, S. C., & Lin, F. L. (2016). Inquiry-based mathematics curriculum design for young children-teaching experiment and reflection. *Eurasia journal of mathematics, science and technology education*, *12*(4), 843-860.

Xu, Z., Chen, Z., Eutsler, L., Geng, Z., & Kogut, A. (2020). A scoping review of digital game-based technology on English language learning. *Educational Technology Research and Development*, *68*(3), 877-904.

Yang, Y. T. C., & Chang, C. H. (2013). Empowering students through digital game authorship: Enhancing concentration, critical thinking, and academic achievement. *Computers & Education*, *68*, 334-344.

Yang, Y.-T. C., Chuang, Y.-C., Li, L.-Y., & Tseng, S.-S. (2013). A blended learning environment for individualized English listening and speaking integrating critical thinking. *Computers & Education, 63*, 285–305.

Yeh, Y. C. (2003). The manual of critical thinking test, level I. *Taipei: Psychological Publishing*.

Yeh, Y. C. (2009). Integrating e-learning into the Direct-instruction Model to enhance the effectiveness of critical-thinking instruction. *Instructional Science*, *37*(2), 185-203.

Yeh, Y. C. (2012). A co-creation blended KM model for cultivating critical-thinking skills. *Computers & Education*, *59*(4), 1317-1327.

Yoo, M. S., & Park, H. R. (2015). Effects of case-based learning on communication skills, problem-solving ability, and learning motivation in nursing students. *Nursing & health sciences*, *17*(2), 166-172.

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., ... & Yukhymenko, M. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of educational research*, *82*(1), 61-89.

Yu, F. Y., & Yang, Y. T. (2014). To See or Not to See: Effects of Online Access to Peer-Generated Questions on Performance. *Educational Technology & Society*, *17*(3), 27-39.

Yunus, A. S. M., Abu, R., Nor, S. M., Tarmizi, R. A., Bakar, K. A., Ali, W. Z. W., ... & Ismail, H. (2005). Generic skills of Malaysian university students. *Bulletin of Higher Education Research. Instititut Penyelidikan Pendidikan Tinggi Negara*, (6), 5-6.

Yusuf, N. and Shah, P., 2018. TEACHERS' PERCEPTION ON BARRIERS OF TEACHING CRITICAL THINKING IN ESL CLASSROOM: A SURVEY. In: *12th International Conference on Language, Education and Innovation (ICLEI)*. [online] pp.66-80. Available at: <http://icsai.org/procarch/12iclei/12iclei-026.pdf> [Accessed 1 February 2021].

Zabit, M. N. M. (2010). Problem-Based Learning on Students' Critical Thinking Skills in Teaching Business Education in Malaysia: A Literature Review. *American Journal of Business Education*, *3*(6), 19-32.

Zainal Shah, N. (2011). *Critical thinking and employability of computer-related graduates: The Malaysian context* (Doctoral dissertation, Dublin City University).

Zhao, B. (2012). Monopoly, economic efficiency and unemployment. *Economic Modelling*, *29*(3), 586-600.

Zohar, A., & Cohen, A. (2016). Large scale implementation of higher order thinking (HOT) in civic education: The interplay of policy, politics, pedagogical leadership and detailed pedagogical planning. *Thinking Skills and Creativity*, *21*, 85–96.

Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, *38*(9), 25-32.

Zagal, J. P., & Altizer, R. (2015, October). Placeholder Content in Game Development: Benefits and Challenges. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (pp. 745-750).

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Al-Nujaifi, M. H., Ow, S. H., Loh, S. C., & Thong, M. (2017). Impact of Digital Game-Based Learning on Mathematical Learning of Elementary School Students: A Mini Review. *2nd International Postgraduate Conference on Applied Science and Technology (IPCAST 2016)*, Langkawi, Malaysia. 173—176.

Al-Nujaifi, M. H., Ow, S. H. (2018). Improving critical thinking skills of elementary school pupils using the ICMDCR framework. *The Int. Conf. Educ.*, Kuala Lumpur, Malaysia. 1—9.

Hussein, M. H., Ow, S. H., Cheong, L. S., Thong, M. K., & Ale Ebrahim, N. (2019). Effects of Digital Game-Based Learning on Elementary Science Learning: A Systematic Review. *IEEE Access*, *7*, 62465–62478. https://doi.org/10.1109/ACCESS.2019.2916324

Hussein, M. H., Ow, S. H., Cheong, L. S., & Thong, M. K. (2019). A Digital Game-Based Learning Method to Improve Students' Critical Thinking Skills in Elementary Science. *IEEE Access*, *7*, 96309–96318. https://doi.org/10.1109/ACCESS.2019.2929089

Hussein, M. H., Ow, S. H., Ibrahim, I., & Mahmoud, M. A. (2020). Measuring instructors continued intention to reuse Google Classroom in Iraq: a mixed-method study during COVID-19. *Interactive Technology and Smart Education*. https://doi.org/10.1108/ITSE-06-2020-0095

Hussein, M. H., Ow, S. H., Elaish, M. M., Jensen, E. O.,. Digital Game-Based Learning in K-12 Mathematics Education: A Systematic Literature Review, *Education and Information Technologies,* 1-33. https://doi.org/10.1007/s10639-021-10721-x